

DEVELOPMENT OF A NOVEL SCREENING METHOD FOR THE  
ISOLATION OF PRECURSOR RNA PROCESSING MUTANTS  
OF *SACCHAROMYCES CEREVISIAE*.

John Morran

A thesis presented for the degree of  
Doctor of Philosophy  
University of Edinburgh

June, 1990



For my family,

In the hope that education may one day be based on the equality of opportunity rather than the ability to pay.

## DECLARATION

I hereby declare that I alone have composed this thesis and that, except where stated, the work presented herein is my own.

JOHN MORRAN

June 1990

## ABSTRACT

The availability of conditional-lethal yeast (*Saccharomyces cerevisiae*) mutants that are defective in the process of pre-mRNA splicing (precursor RNA processing; *prp*) has greatly facilitated the characterisation of components of the splicing machinery. When this project was initiated nine *prp* complementation groups had been defined (*prp2-prp11*), all of which accumulate pre-mRNA at the expense of mRNA when incubated at the non-permissive temperature. Pre-mRNA splicing is a complex and dynamic process relying on many gene products for its completion and therefore many more complementation groups remained to be identified. Determination of a *prp* phenotype has relied on the direct measurement of protein to RNA ratios and on the Northern blot analysis of conditional-lethal mutants incubated at the restrictive temperature. Such analyses are both time-consuming and labour-intensive. In this thesis I present the development of a novel screening procedure which positively identifies *prp* mutants. I have fused a yeast intron-containing gene to the *lacZ* gene of *E.coli* such that only the pre-mRNA generated from this fusion can encode an active  $\beta$ -galactosidase fusion-protein. This gene-fusion has been introduced into a *prp2* strain and the encoded pre-mRNA has been shown to accumulate on incubation at the non-permissive temperature. This pre-mRNA accumulation results in an increase in  $\beta$ -galactosidase activity. Exploiting this observation a simple plate assay has been developed and used to screen a pool of temperature-sensitive mutants for defects in pre-mRNA splicing. A number of *prp* mutants have been identified and I present the results of their initial characterisation.



## CONTENTS

	Page no.
Declaration	i
Abstract	ii
Contents	iii
Acknowledgements	xii
Abbreviations	xiii
<u>CHAPTER 1</u>	
INTRODUCTION	1
1.1 RNA SPLICING	1
1.2 INTRONS IN NUCLEAR PROTEIN-ENCODING GENES	2
1.3 MECHANISM OF PRE-mRNA SPLICING	2
1.4 SEQUENCE REQUIREMENTS FOR PRE-mRNA SPLICING	4
1.4.1 Mammalian Pre-mRNA Transcripts	4
1.4.2 Yeast Pre-mRNA Transcripts	8
1.4.3 Comparison of Mammalian and Yeast Sequence Requirements	13
1.5 MAMMALIAN SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES	14
1.5.1 Introduction	14
1.5.2 Evidence for the Involvement of snRNPs in Splicing	15
1.5.3 snRNP Proteins	19
1.6 OTHER FACTORS INVOLVED IN MAMMALIAN SPLICING	21
1.7 MAMMALIAN SPLICING COMPLEXES	22
1.7.1 Isolation and Identification	22
1.7.2 Temporal Order of snRNP Addition in Spliceosome Assembly	23
1.8 YEAST SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES	25

1.8.1	Introduction	25
1.8.2	Evidence for the Involvement of Yeast snRNPs in Splicing	26
1.8.3	Yeast snRNP Proteins	28
1.9	GENETIC AND BIOCHEMICAL APPROACHES TO THE ISOLATION OF FACTORS INVOLVED IN PRE-mRNA SPLICING IN YEAST	29
1.10	YEAST SPLICING COMPLEXES	32
1.10.1	Isolation and Identification	32
1.10.2	Temporal Order of snRNP Interaction in Spliceosome Assembly	33
1.11	PRE-mRNA SPLICING AND THE CONTROL OF GENE EXPRESSION	34
1.11.1	Alternative Splicing	34
1.11.2	Introns: Effects on Gene Expression	37
1.12	INTRONS AND EVOLUTION	39
1.13	THIS THESIS	41
	 <u>CHAPTER 2</u>	 42
	MATERIAL AND METHODS	42
2.1	MATERIALS	42
2.1.1	Suppliers of Laboratory Reagents	42
2.1.2	Growth Media	42
2.1.2.1	Bacterial Media	43
2.1.2.2	Yeast Media	43
2.1.3	Bacterial Strains	44
2.1.4	Yeast Strains	44

TABLE 2.1	BACTERIAL STRAINS	45
TABLE 2.2	<i>SACCHAROMYCES CEREVISIAE</i> STRAINS	46
2.1.5	General Plasmid and Bacteriophage Vectors	52
2.1.6	Synthetic Oligodeoxynucleotides	52
TABLE 2.3	GENERAL PLASMID AND BACTERIOPHAGE VECTORS	53
TABLE 2.4	SYNTHETIC OLIGODEOXYNUCLEOTIDES	54
2.2	METHODS	56
2.2.1	General Guidelines	56
2.2.1.1	Centrifugation	56
2.2.1.2	Sterilization of Solutions	56
2.2.1.3	Deionisation of Solutions	56
2.2.1.4	Autoradiography	57
2.2.1.5	Photography	57
2.2.2	Bacteria and Phage: General Microbial Methods	57
2.2.2.1	Propagation and Storage	57
2.2.2.2	Preparation of Competent Cells	57
2.2.2.3	Transformation	58
2.2.2.4	Isolation of M13 Plaques	58
2.2.3	Yeast: General Microbiological and Genetic Methods	59
2.2.3.1	Propagation and Storage	59
2.2.3.2	Generation of Diploid Strains	59

2.2.3.3	Sporulation and Spore Recovery	59
2.2.3.4	Mating-Type Analysis	60
2.2.3.5	Complementation Analysis	60
2.2.3.6	Transformation	60
2.2.3.7	UV Mutagenesis and Isolation of $r^S$ Colonies	61
2.2.3.8	UV Mutagenesis: Survival Curve	62
2.2.4	General Nucleic Acid Methods	63
2.2.4.1	Storage	63
2.2.4.2	Quantitation of Nucleic Acid	63
2.2.4.3	Precipitation with Ethanol	63
2.2.4.4	Extraction with Phenol/Chloroform	63
2.2.5	DNA Methods	64
2.2.5.1	Plasmid DNA Isolation: Small Scale Preparations	64
2.2.5.2	Plasmid DNA Isolation: Midi-Preparations	65
2.2.5.3	Plasmid DNA Isolation: Large Scale Preparations	66
2.2.5.4	Preparation of M13 RF DNA	66
2.2.5.5	Restriction Enzyme Digestion of DNA	67
2.2.5.6	Agarose Gel Electrophoresis	67
2.2.5.7	Isolation of Restricted DNA Fragments	67
2.2.5.8	Ligation of DNA	68
2.2.5.9	Filling in DNA Termini with Recessed 3' Ends	69
2.2.5.1	Labelling DNA by the Random Priming Method	69
2.2.5.11	Measurement of Radioactivity Incorporated into DNA	70
2.2.5.12	5' End Labelling of Oligodeoxynucleotides	70
2.2.5.13	Purification of Labelled Oligodeoxynucleotides	71
2.2.5.14	Colony Blotting and Hybridisation	71

2.2.5.15 M13 Plaque Hybridisation Using DNA Fragment Probes	72
2.2.5.16 Plaque Hybridisation Using Oligodeoxynucleotide Probes	73
2.2.6 Site-Directed Mutagenesis	74
2.2.6.1 5' Phosphorylation of Mutagenic Oligodeoxynucleotides	74
2.2.6.2 Annealing Conditions	74
2.2.6.3 Site-Directed Mutagenesis: Annealing and Extension	76
2.2.6.4 Transformation	76
2.2.7 DNA Sequencing	77
2.2.7.1 Preparation of Template DNA	77
2.2.7.2 Sequencing Solutions	78
2.2.7.3 Sequencing Reactions	78
 TABLE 2.5 COMPOSITION OF dNTP SOLUTIONS	 80
 TABLE 2.6 COMPOSITION OF ddNTP SOLUTIONS	 81
2.2.7.4 Sequencing Gels	82
2.2.8 RNA Methods	83
2.2.8.1 Protection of RNA from RNases	83
2.2.8.2 Extraction of RNA from Yeast	84
2.2.8.3 RNA Agarose Gel Electrophoresis	84
2.2.8.4 Northern Blotting and Hybridisation	84
2.2.8.5 Primer Extension	85
2.2.9 $\beta$ -Galactosidase Assays	86
2.2.9.1 Induction of the <i>GAL1</i> Promoter	86
2.2.9.2 $\beta$ -Galactosidase Assays	87

<u>CHAPTER 3</u>	89
CONSTRUCTION OF PLASMIDS USED IN MUTANT SCREENING AND IN THE ANALYSIS OF PRE-mRNA COMMITMENT TO SPLICING	
3.1 INTRODUCTION	89
3.2 RATIONALE FOR CONSTRUCTION OF THE SCREENING PLASMIDS	89
3.3 PLASMID CONSTRUCTION	91
3.3.1 Parent Plasmids	92
3.3.1.1 pBEMAT/S	92
3.3.1.2 pDEV19	92
3.3.1.3 pSPlacZ	93
3.3.1.4 pBM272 $\Delta$ RI	94
3.3.2 Steps in Plasmid Construction	94
3.3.3 Site-Directed Mutagenesis	95
3.3.3.1 Mutant Phage Detection	95
3.3.3.2 Changes to <i>MATa1</i> Region in M13mp9	96
3.4 Summary	99
 <u>CHAPTER 4</u>	 100
FEASIBILITY AND DEVELOPMENT OF THE SCREENING PROCEDURE	
4.1 THE FEASIBILITY TO A <i>prp</i> MUTANT SCREEN	100
4.1.1 Primer Extension Analysis of the <i>MATa1-lacZ</i> Fusion-Transcript from pJBM-4 in Wild-Type and <i>prp</i> Strains	101
4.1.2 Analysis of Plasmid Encoded $\beta$ -galactosidase Activity	103

4.1.2.1	DBY745	103
4.1.2.2	DJY36 ( <i>prp2</i> )	105
4.1.2.3	SPJ8.31 ( <i>prp8</i> )	106
4.1.2.4	DJY72 ( <i>prp1</i> )	108
4.1.3	Summary of Results Indicating Screen Feasibility	109
4.2	DEVELOPMENT OF THE SCREENING PROCEDURE	110
4.2.1	Choice of Wild-Type Strain	110
4.2.1.1	Strains Producing "Petite" Mutants	110
4.2.1.2	Glucose Repressibility	111
4.2.1.3	Galactose Induction Mutants	111
4.2.2	Growth Media	112
4.2.3	Detection of $\beta$ -galactosidase Activity	112
	<u>CHAPTER 5</u>	116
	ISOLATION AND CHARACTERISATION OF POTENTIAL CONDITIONAL-LETHAL <i>prp</i> MUTANTS	
5.1	GENERATION OF TEMPERATURE-SENSITIVE MUTANTS	117
5.2	SCREENING FOR POTENTIAL <i>prp</i> MUTANTS	117
5.2.1	DBY746	118
5.2.2	DBY745	118
5.3	NORTHERN BLOT ANALYSIS OF POTENTIAL <i>prp</i> MUTANTS	120
5.3.1	<i>t</i> <sup>s</sup> Mutants Showing a Classic <i>prp</i> Northern Phenotype	122

5.3.2	Mutants Which May Have Been Defective in Splicing	122
5.3.2.1	20a	123
5.3.2.2	36a	123
5.3.2.3	109i	125
5.3.2.4	121e	125
5.3.2.5	196a	126
5.3.3	Mutants in Which There Appears to be a Small Decrease in Splicing Efficiency	126
5.3.3	Mutants With Altered Transcript Size	127
5.4	COMPLEMENTATION ANALYSIS OF POTENTIAL <i>prp</i> MUTANTS	128
5.4.1	Northern Analysis of Strains Crossed to <i>prp25</i> and <i>prp26</i>	129
5.5	<i>t<sup>s</sup></i> PROFILE OF ISOLATED MUTANTS	131
5.6	DISCUSSION	133
5.6.1	Comparison of the Efficiency of the Screening Procedure Using DBY745 and DBY746	133
5.6.2	Comparison of the Plate-Screening Method with Other Screening Procedures	134
5.6.2.1	Advantages of the Plate-Screening Procedure	134
5.6.2.2	Disadvantages of the Plate-Screening Procedure	135
5.6.3	New Alleles of Pre-Existing <i>prp</i> Mutants Isolated by the Plate-Assay Screening Procedure	136
5.6.4	Other Types of Mutants that Could be Isolated Using the Plate-Assay Screen	137



5.6.5	Future Work	138
5.6.6	Other Possible Screens	138
<u>REFERENCES</u>		140
APPENDIX: COMMITMENT OF A PRE-mRNA TRANSCRIPT TO		
	SPLICING	157
INTRODUCTION		157
ANALYSIS OF PRE-mRNA COMMITMENT		
	UTILISING pJBM-1 TO -5	158
DISCUSSION		163

## ACKNOWLEDGEMENTS

I would like to thank all the people (past and present) of the Beggs' lab. for making my time in Edinburgh enjoyable. Thanks to Gordon, Mike, Derek, Jean, Dave, Frankie, Susan, Helen, Mary (hairy!), Jeremy, Nik and Claire. Special thanks to Margaret and Sarah for help with the proof-reading and for putting up with the various noises emitted whilst sitting at the word-processor. To Derek and Lizzie for help with the graphs and to Janet for editing and typing. Erica, my "wee" twin, I'm glad it was you I shared the full term with....., thanks for all the dirty jokes! Thanks to David Finnegan and John Scaife for their interest in my project and their advice. To Donne, for putting up with me when it was bad and to Richie and Karen for the smiles and the use of their T.V. during the World Cup (though they don't know it yet!). Last but not least a big thank-you to my three tremendous flat-mates; Alison (thanks for the proof-reading big Chalkie!), Hazel and Joan. Staying in the flat was undoubtedly the best part of three and a half very long years.

My continuation in higher education could not have been achieved without the love and support of my family and I cannot begin to express my gratitude to them.....

This work was supported by a three-year MRC postgraduate studentship.

## ABBREVIATIONS

ATP	:	adenosine 5' -triphosphate
bp	:	base pair
BSA	:	bovine serum albumin
Ci	:	Curie ( $2.2 \times 10^{12}$ dpm)
cms	:	centimetres
CTP	:	cytosine 5' -triphosphate
D	:	Dalton
dATP	:	2' -deoxyadenosine 5' -triphosphate
dCTP	:	2' -deoxycytosine 5' -triphosphate
DEP	:	diethylpyrocarbonate
dGTP	:	2' -deoxyguanosine 5' -triphosphate
dNTPs	:	deoxyribonucleotides
dTTP	:	2' -deoxythymidine 5' -triphosphate
ddATP	:	2', 3' -dideoxyadenosine 5' -triphosphate
ddCTP	:	2', 3' -dideoxycytosine 5' -triphosphate
ddGTP	:	2', 3' -dideoxyguanosine 5' -triphosphate
ddTTP	:	2', 3' -dideoxythymidine 5' -triphosphate
dH <sub>2</sub> O	:	distilled water
DMSO	:	dimethyl sulphoxide
DNA	:	deoxyribonucleic acid
DNase	:	deoxyribonuclease
DTT	:	dithiothreitol
E1	:	exon one
E2	:	exon two
EDTA	:	ethylene-diamine-tetracetic acid
g	:	acceleration due to gravity ( $9.81\text{m.s}^{-2}$ )

g	:	gram
GTP	:	guanosine 5' -triphosphate
HEPES	:	4-(2-hydroxyethyl)-1-piperazine-ethanesulphonic acid
hnRNP	:	heterogeneous nuclear RNP
IPTG	:	isopropyl $\beta$ -D-thiogalactoside
IVS	:	intervening sequence
IVS-E2	:	intervening sequence-exon two lariat intermediate
k	:	kilo (ie. $\times 10^3$ )
l	:	litre
ml	:	millilitre (ie. $\times 10^{-3}$ )
m	:	metre
M	:	molar (moles/litre)
MOPS	:	4-morpholine-propane-sulphonic acid
mRNA	:	messenger RNA
m <sup>3</sup> G	:	2, 2, 7-trimethylguansine
MW	:	molecular weight
n	:	nano (ie. $\times 10^{-9}$ )
nts	:	nucleotides
OD <sub>600nm</sub>	:	optical density with respect to light of wavelength 600nm
ONPG	:	0-nitrophenyl $\beta$ -D galactopyranoside
PEG	:	polythylene glycol
pers.comm.	:	personal communication
pmol	:	picomole (ie. $\times 10^{-12}$ )
PMSF	:	phenyl-methyl-sulphonyl fluoride
<i>prp</i>	:	precursor RNA processing
pre-mRNA	:	precursor messenger RNA
psi	:	pounds per square inch
Pu	:	purine nucleotide

Py	:	pyrimidine nucleotide
RF	:	replicative form
RNA	:	ribonucleic acid
RNase	:	ribonuclease
RNP	:	ribonucleoprotein
rpm	:	revolutions per minute
rRNA	:	ribosomal RNA
SDS	:	sodium dodecyl sulphate
SDM	:	site-directed mutagenesis
secs	:	seconds
snRNA	:	small nuclear RNA
snRNP	:	small nuclear RNP
TCA	:	trichloroacetic acid
TEMED	:	N, N, N', N'-tetramethyl-ethylenediamine
t <sup>s</sup>	:	temperature-sensitive
Tris	:	2-amino-2-hydroxymethyl-propane-1, 3-diol
tRNA	:	transfer RNA
UTP	:	uridine 5' -triphosphate
UV	:	ultra-violet light
V	:	Volt
v/v	:	volume per unit volume
W	:	Watt
w/o	:	without
w/v	:	weight per unit volume
X-gal	:	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside
$\mu$	:	micro (ie. $\times 10^{-6}$ )
$\mu\text{g}$	:	microgram
$\mu\text{l}$	:	microlitre

## CHAPTER 1

### INTRODUCTION

#### 1.1 RNA SPLICING

In the mid 1970s it was discovered that the coding regions of many eukaryotic genes are interrupted by stretches of DNA which do not encode RNA found in the mature transcript (see Crick, 1978). It was subsequently found that these intervening sequences (introns; see Gilbert, 1978) are removed from a primary RNA transcript by a process now known as pre-mRNA splicing. Introns are divided into 4 classes based on the mechanism of their removal, and the conserved sequence and structural elements residing in them. These 4 classes are;

- (i) those found in tRNA genes (Abelson, 1979),
  - (ii) Group I introns, first described in *Tetrahymena* rRNA genes but now also found in many fungal mitochondrial genes, chloroplast genes and in some bacteriophage T4 genes (Cech and Bass, 1986),
  - (iii) Group II introns, found in fungal mitochondrial and chloroplast genes (Cech and Bass, 1986),
  - (iv) those found in nuclear protein-encoding genes (Green, 1986; Padgett *et al.*, 1986).
- Recently a Group IV intron has been found in the U6 snRNA gene of *Schizosaccharomyces pombe* (Tani and Ohshima, 1989).

This thesis is concerned with the removal of introns from protein-encoding genes of eukaryotic organisms and therefore this Chapter will introduce relevant aspects of this process. In this thesis *Saccharomyces cerevisiae* and yeast, and TACTAAC and UACUAAC, will be used interchangeably.

## 1.2 INTRONS IN NUCLEAR PROTEIN-ENCODING GENES

In higher eukaryotic organisms protein-encoding genes can contain multiple introns ranging in size from 31 nucleotides to 100kb (Green, 1986; Padgett *et al.*, 1986; Smith *et al.*, 1989a). This arrangement facilitates the process of alternative splicing, thereby generating diverse combinations of exons (RNA sequences present in the mature transcript) (see Section 1.11.1). Alternative splicing is extremely important in the control of gene expression (Leff and Rosenfeld, 1986; Breibart *et al.*, 1987; Smith *et al.*, 1989a).

In protein-encoding genes of yeast it appears that the presence of introns is the exception rather than the rule (Teem *et al.*, 1984). In contrast to mammals the introns present in yeast genes fall into a distinct size range of 300-500 nucleotides. There is usually only one intron present in any given gene and it is generally situated at the 5' end of the primary transcript. To date only one exception to these observations has been found; the *MATa1* gene contains two small introns of 54 and 52 nucleotides (Miller, 1984), however no biological significance for this arrangement has yet been observed (Kohrer and Domdey, 1988; Ner and Smith, 1989).

## 1.3 MECHANISM OF PRE-mRNA SPLICING

Unravelling the biochemical mechanism of splicing required the development of an efficient *in vitro* splicing system which could faithfully reproduce the splicing pattern observed *in vivo*. Early attempts at such a development relied on systems which coupled transcription and splicing in whole cell or nuclear extracts (Weingartner and Keller, 1981; Kole and Weissman, 1982; Padgett *et al.*, 1983a); or relied on the addition of exogenous polyadenylated RNA precursors to the splicing extract (Goldenberg and Raskas, 1981; Hernandez and Keller, 1983; Goldenberg and Hauser, 1983). Such strategies either produced splicing of a low efficiency or involved manipulations that were extremely time-consuming.

The biochemical pathway of splicing was eventually dissected utilising HeLa cell nuclear

## FIGURE 1.1: INTRON STRUCTURE

The conserved sequence elements of both mammalian and yeast introns are indicated. Vertical bars indicate the 5' and 3' splice sites. The numbers above each nucleotide refers to their position with respect to the discussion of their mutational analyses (see Section 1.4). The asterisk indicates the branchpoint nucleotide in yeast introns.

## FIGURE 1.2: MECHANISM OF SPLICING

The mechanism of nuclear pre-mRNA splicing in mammalian and yeast cells.



Figure 1.1

MAMMALIAN



S.CEREVISIAE

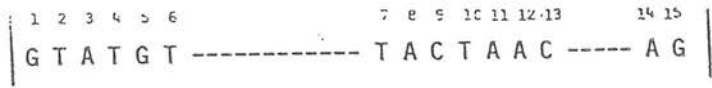
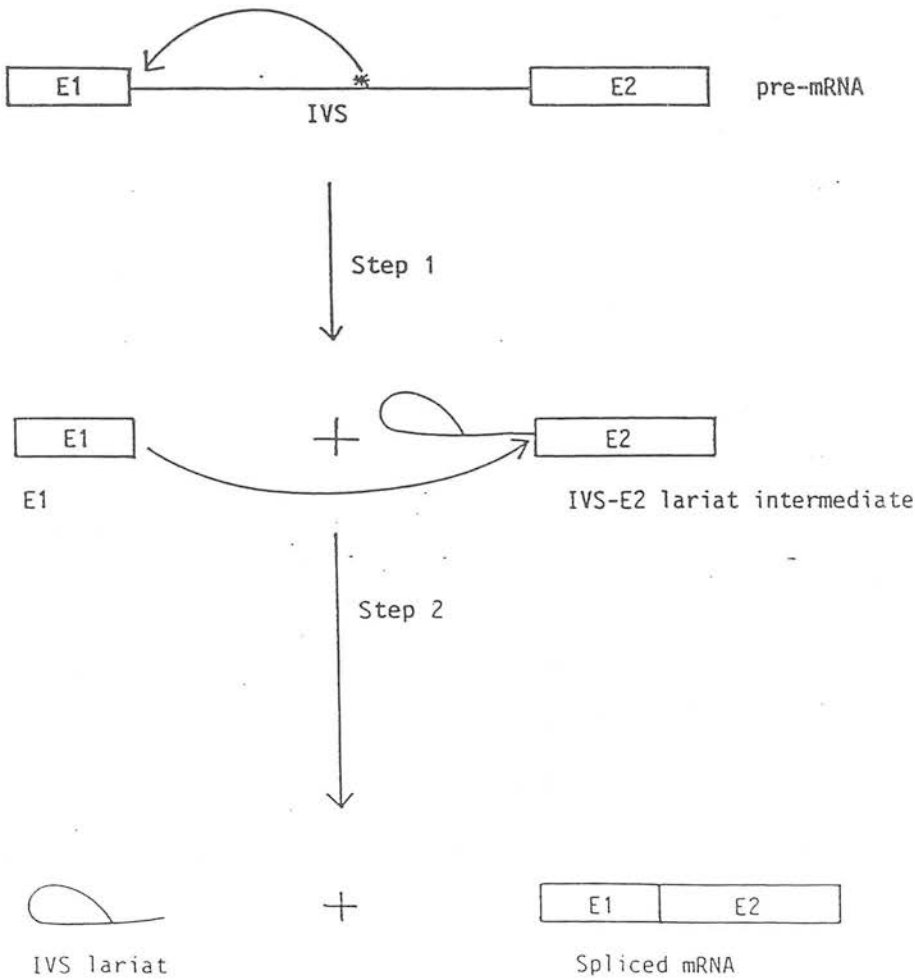


Figure 1.2



extracts, which could efficiently and faithfully splice exogenous pre-mRNA transcripts generated *in vitro* (Krainer *et al.*, 1984; Hardy *et al.*, 1984; Ruskin *et al.*, 1984; Grabowski *et al.*, 1984; Padgett *et al.*, 1984; Keller, 1984; Konarska *et al.*, 1985; Padgett *et al.*, 1985). These studies led to the proposal of a two-step pathway (see Fig. 1.2). After a characteristic lag period, usually 30-45 minutes, the first chemical modification to the pre-mRNA can be detected. Step 1 of the splicing reaction proceeds with the cleavage of the transcript at the 5' splice site producing exon 1 (E1) and intron lariat-exon 2 (IVS-E2) intermediates. The 5' end of the intron is joined via a 5'-2' phosphodiester bond to an A nucleotide usually situated 18-40 nucleotides upstream of the 3' splice site. Step 2 involves the cleavage of the IVS-E2 at the 3' splice site and the ligation of E1 to E2 to form the mature RNA transcript. The intron (IVS) is recovered as a lariat structure.

It is not known whether bond cleavage and formation at each step in the reaction occur concomitantly. There is no net change in the number of phosphodiester bonds and therefore theoretically these reactions require no phosphate donor i.e. no input of energy. Because of this and the similarity to the excision of the self-splicing introns of Groups I and II (Cech and Bass, 1986), it is considered that bond breaking and formation are the result of a series of nucleophilic attacks as indicated in Fig. 1.2 and at each step occur concomitantly. However splicing *in vitro* does require ATP and this requirement may be for splicing-complex assembly and conformational changes within this assembly (e.g. see Cheng and Abelson, 1987). In yeast cell extracts, splicing proceeds by an identical mechanism (Lin *et al.*, 1985).

For the analysis of splicing *in vitro* to be of any benefit the same mechanism of splicing must take place *in vivo* and thus the same molecules must be identifiable. HeLa cell nuclear extracts (but not cytoplasmic extracts) contain a significant number of RNA molecules that harbour 2'-5' phosphodiester branches (Wallace and Edmonds, 1983). The nucleotide at the branch site of these molecules was found to be nearly always an A, though C and to a lesser extent U branched nucleotides were also found. Subsequently the intermediates of the splicing reaction were identified *in vivo* in mammalian cells (Zeitlin and Efstratiadis, 1984) and in yeast cells (Rodriguez *et al.*, 1984; Domdey *et al.*, 1984).

The presence of a m<sup>7</sup>G(5')ppp(5')N cap structure at the 5' end of intron-containing transcripts is not required for splicing in mammalian nuclear extracts, though when analysed *in vitro* its presence does increase both the efficiency of splicing- complex formation and fidelity of the splicing reaction (Konarska *et al.*, 1984; Patzelt *et al.*, 1987; Inoue *et al.*, 1989). In yeast extracts the presence of the cap structure on the transcript appears to have no effect on splicing (Lin *et al.*, 1985). Polyadenylation or correct 3' end formation of pre-mRNA transcripts are not required for splicing (Krainer *et al.*, 1984).

Since nuclear pre-mRNA introns present in both mammalian and yeast transcripts are spliced by the same mechanism a search was undertaken for conserved sequences within the introns of both species (and indeed in other species) (Mount, 1982; Keller and Noon, 1984; Teem *et al.*, 1984). This search pointed to the conservation of sequences at the 5' end, branchpoint and 3' end of the introns. To determine the importance of these conserved sequences in the process of pre-mRNA splicing, numerous experiments were performed involving deletion and mutational analysis of these elements. The effects of the changes introduced have been analysed extensively both *in vivo* and *in vitro*, and generally the observations made using both approaches have been consistent. However the effects of the mutations appear to be more pronounced when analysed *in vitro*.

## 1.4 SEQUENCE REQUIREMENTS FOR PRE-mRNA SPLICING

### 1.4.1 Mammalian Pre-mRNA Transcripts

Naturally occurring mutations in the 5' splice site of globin gene introns not only reduce the efficiency of splicing but also activate cryptic 5' splice sites *in vitro* and *in vivo* (Felber *et al.*, 1982; Treisman *et al.*, 1982, 1983; Wieringa *et al.*, 1983; Krainer *et al.*, 1984). Mutations at G1 or T2 (refer to Fig.1.1) produce the most pronounced effect on splicing efficiency, however there are contrasting reports regarding the qualitative effects of such

mutations. In some reports 5' cleavage is completely inhibited (Treisman *et al.*, 1982, 1983; Wieringa *et al.*, 1983; Green *et al.*, 1983) whereas in others the efficiency of step 1 is reduced and step 2 is inhibited (Aebi *et al.*, 1986, 1987). Mutations in the conserved GT nucleotides not only reduce the efficiency of active spliceosome formation but also cause the accumulation of a complex containing the intermediates of the splicing reaction (Lamond *et al.*, 1987). This observation suggests that mutations in the conserved GT nucleotides cause a block in step 2 of the splicing reaction. Other mutations in the 5' splice site region, though exhibiting less of an effect on the efficiency of splicing, lead to the activation of cryptic splice sites (Treisman *et al.*, 1983) and "exon-skipping" (Aebi *et al.*, 1986). The 5' splice site consensus sequence (indicated in Fig. 1.1) is the sequence used most efficiently in the splicing process (Aebi *et al.*, 1986, 1987; Lear *et al.*, 1990). It is considered that this is a consequence of its optimum complementarity to the 5' end of the U1 snRNA (Aebi *et al.*, 1987). However it appears that this complementarity to the U1 snRNA is not the only determinant of 5' splice site choice. The context in which the 5' splice site resides appears to influence its relative use (Nelson and Green, 1988). Recently it has been postulated that the sequence at the 5' splice site may be affected by the length of the intron (Fields, 1990).

The mammalian branch-acceptor region is not as highly conserved as that for yeast transcripts. Deletion of the authentic branchpoint region results in the use of cryptic branch-acceptor sites (Reed and Maniatis, 1985), indeed cryptic sites need not exhibit any sequence similarity to the branchpoint consensus (Padgett *et al.*, 1985). Utilisation of cryptic branchpoints appears to decrease the efficiency of splicing (Padgett *et al.*, 1985). Branchpoints are usually situated 18-40 nucleotides upstream of the 3' splice site (Ruskin *et al.*, 1984; Reed, 1989) and the minimum distance from the 5' splice site that allows efficient splicing is ~50 nucleotides (see Smith *et al.*, 1989a). Increasing the distance separating the 3' splice site from the branchpoint above ~70 nucleotides decreases the efficiency of step 2 of the splicing reaction (Reed, 1989). Although most branches are formed at an A nucleotide any nucleotide can act as a branch-acceptor, however G or U nucleotides form IVS-E2 molecules incapable of

undergoing step 2 of the splicing reaction (Hornig *et al.*, 1986). Coupled with the results of mutations at positions G1 and T2 of the 5' splice site, these observations suggest that specific nucleotides are required on either side of the phosphodiester bond formed at the branchpoint to enable step 2 of the splicing reaction to take place. Mutational analysis of sequences surrounding the branchpoint have indicated that this region affects the efficiency of branchpoint recognition and utilisation (Reed and Maniatis, 1988). Thus the weak consensus derived for the branchpoint region is important in pre-mRNA splicing. Some mammalian branchpoints are situated well upstream of the 3' splice site i.e. ~100 nucleotides (Reed, 1989; Goux-Pelletan *et al.*, 1990). It has been suggested that such sites are utilised in the process of alternative splicing (Reed, 1989). Such sites are normally followed by an extremely long (Py)<sub>n</sub> tract, and such long tracts do not require an AG (3' cleavage site) in close proximity to enable efficient completion of step 1 of the splicing reaction (Reed, 1989).

There are two elements present at the 3' splice site region of introns which are of importance in the splicing process. The 3' cleavage site immediately follows the highly conserved AG dinucleotide which is itself preceded by a (Py)<sub>n</sub> tract (n>10) (See Fig. 1.1). Deletion of the (Py)<sub>n</sub> tract from the first intron of the human  $\beta$ -globin gene and the major late transcript of adenovirus 2 completely abolishes lariat formation (Frendewey and Keller, 1985; Reed and Maniatis, 1985a; Ruskin and Green, 1985c) and spliceosome assembly (Frendewey and Keller, 1985; Bindereif and Green, 1986). The efficiency of splicing appears to decrease sequentially with the progressive reduction in size of the (Py)<sub>n</sub> tract, which can also lead to the utilisation of cryptic 3' splice sites (Wieringa *et al.*, 1984). Mutations of the AG dinucleotide can either moderately inhibit step 1 of splicing (Reed and Maniatis, 1985) or completely abolish it (Aebi *et al.*, 1986). In addition step 2 is completely inhibited by mutations in the AG dinucleotide, indeed cryptic 3' cleavage sites can be activated in response to such mutations (Aebi *et al.*, 1986). Mutations of the AG dinucleotide inhibit spliceosome assembly (Lamond *et al.*, 1987). In *cis*-competition assays R. Reed was able to define more clearly the relationship between the (Py)<sub>n</sub> tract and the AG cleavage site and concluded that there is a sequence and distance requirement for efficient completion of steps 1 and 2 of splicing (Reed, 1989). From

this analysis it appears that the (Py)<sub>n</sub> tract is required primarily for step 1 and the AG dinucleotide is required (though not essential) for efficient completion of step 2. Recently it has been observed that the branchpoint is specified by both its sequence context and by the adjacent downstream (Py)<sub>n</sub> tract (Smith *et al.*, 1989b). Indeed splicing efficiency is greater when the (Py)<sub>n</sub> tract is situated immediately downstream of the branchpoint region (Reed, 1989). Specification of the AG cleavage site, which is normally the first AG dinucleotide downstream of the branchpoint, is proposed to be by a scanning mechanism initiated at the branchpoint region (Smith *et al.*, 1989b). It was previously considered that in mammalian introns the branchpoint was determined, in part, by its distance from the 3' cleavage site, therefore this scanning model rationalises the observation that some branchpoints are located at greater than expected distances upstream of the splice site.

Deletion of intron sequences (other than the conserved elements) appears to have little effect on the efficiency of splicing (Wieringa *et al.*, 1984; van Santen and Spritz, 1985). However in the case of alternative splicing, intron sequences may play a role in splice site selection (Goux-Pelletan *et al.*, 1990; Helfman *et al.*, 1990). Introns below the size of 80 nucleotides are spliced inefficiently from their transcripts (Wieringa *et al.*, 1984), however introns of significantly smaller sizes have been identified though this may reflect the splicing process in individual organisms (Karn *et al.*, 1983; Kay *et al.*, 1987).

In contrast to non-conserved intron sequences, exon sequences do influence the efficiency of splicing and indeed the choice of splice sites (Reed and Maniatis, 1986; Lear *et al.*, 1990). Both steps of splicing can be detected when as little as 4 nucleotides are present in exon 2, however the efficiency of splicing increases as the length of exon 2 increases (Turnbull-Ross *et al.*, 1988). Indeed exon 2 is not required for step 1 of splicing (Fredericewey and Keller, 1985). A transcript containing 20 nucleotides in exon 1 can undergo both steps of splicing (Parent *et al.*, 1987).

#### 1.4.2 Yeast Pre-mRNA Transcripts

It is clear that mutations at most positions in the consensus sequences (see Fig. 1.1) affect the efficiency of splicing and primarily they cause an accumulation of pre-mRNA. However they do so to varying degrees and indeed different nucleotide substitutions at the same positions can produce varied effects (Jaquier *et al.*, 1985; Vijaraghavan *et al.*, 1986; Fouser and Friesen, 1986). In different transcripts there also appear to be distinct responses to comparable mutations (see Newman *et al.*, 1985; Vijaraghavan *et al.*, 1986). The effects of nucleotide substitutions and transcript-specific responses appear to be quantitative rather than qualitative.

Mutations at G1 and T2 of the 5' splice site consensus region allow step 1 of the splicing reaction to proceed but result in the partial accumulation of IVS-E2 intermediate molecules which cannot undergo step 2 (Newman *et al.*, 1985; Vijaraghavan *et al.*, 1986; Fouser and Friesen, 1986; Siliciano and Guthrie, 1988). This suggests that 5' splice site cleavage and lariat formation do not require a G nucleotide at position 1 but that step 2 does. This is consistent with the observations made for similar mutations in mammalian introns (see Section 1.4.1). This hypothesis is further supported by the observation that mutations at position G5 produce aberrant 5' cleavage events which result in "dead-end" IVS-E2 molecules. These aberrant sites are situated close to the normal 5' cleavage site (~5 nucleotides away) and cleavage can occur before an A, C or U nucleotide which then can form 5'-2' phosphodiester bonds with the normal branchpoint A nucleotide (Parker and Guthrie, 1985; Jaquier *et al.*, 1985; Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986). Such molecules are unable to undergo step 2 of the splicing reaction. The sites of aberrant 5' cleavage (caused by a mutation at G5) bear no resemblance to the normal sites of 5' cleavage and it has been suggested that 5' splice site recognition and cleavage are distinct processes. This is supported by the observation that mutations of the yeast U1 snRNA which are complementary to those at position G5 of the 5' splice site, not only increase the efficiency of correct 5' splice site cleavage (indicating specific recognition of this site) but also increased the efficiency of the aberrant 5' cleavage event (Siliciano and Guthrie, 1988; Seraphin *et al.*, 1988). These results suggest that the two 5'

cleavage events i.e. correct and aberrant, arise from the one specific U1 snRNA to 5' splice site recognition event, which can then result in one of the two types of cleavages. They also suggest that a specific nucleotide, G5, is involved in the fidelity of 5' cleavage. These aberrant cleavage events are therefore not comparable to the use of cryptic sites observed in mammalian transcripts where the 5' splice sites are probably recognised independently by individual U1 snRNP particles (Chabot and Steitz, 1987b; Aebi *et al.*, 1987; Nelson and Green, 1988; Lear *et al.*, 1990). However in the *MATa1* transcript, two cryptic 5' splice sites located in exon 1 which function *in vitro* and *in vivo* have been identified, though the biological significance of this observation (if any) is unknown (Kohrer and Domdey, 1988).

Consistent with the *in vivo* observations that 5' cleavage is completely blocked, mutations at positions G1 and T2 prevent spliceosome formation *in vitro* (Vijayraghavan *et al.*, 1986). Mutations at T6 appear to have a limited affect on the efficiency of splicing (Pikielny *et al.*, 1983; Pikielny and Rosbash, 1985; Fouser and Friesen, 1986), however they can be "unmasked" by coupling them with mutations in the UACUAAC sequence i.e. the joint consequence of such double mutants is greater than the sum of their individual effects (Seraphin and Rosbash, 1989a). This suggests an interaction between the 5' splice site and the UACUAAC branchpoint sequence. Mutations at position T4 have no measurable effect on the efficiency of splicing though cells harbouring such mutations exhibit reduced rates of growth (Siliciano and Guthrie, 1988; Seraphin and Rosbash, 1989a). Mutations at this position may affect the commitment of pre-mRNA to splicing and thus explain the reduced growth rate of cells harbouring essential genes mutated at this position (see the Appendix and Legrain and Rosbash, 1989).

Deletion of the TACTAAC sequence results in the complete inhibition of splicing *in vivo* (Langford and Gallwitz, 1983; Pikielny and Rosbash, 1985). Utilising the actin gene mutations at positions T7 and A8 of the transcript (see Fig. 1.1) appear to have little effect on the efficiency of splicing though an accumulation of pre-mRNA and a decrease in the level of mRNA can be observed (Jaquier *et al.*, 1985; Fouser and Friesen, 1986). Conflicting results



indicate that the same mutations at position A8 do have a pronounced effect on the efficiency of splicing (Langford and Gallwitz, 1983). Such differences probably reflect the varied sensitivities of the assays used. Mutations at position A8 on the *RP51A* transcript produce a considerable increase in the level of pre-mRNA. This effect is enhanced by coupling the A8 mutation with a mutation at T6 of the 5' splice site (as indicated above) (Seraphin and Rosbash, 1989a). Mutations at position C9 appear to have a limited effect on the efficiency of splicing (Langford and Gallwitz, 1983; Jaquier *et al.*, 1985). However contradictory results have been obtained indicating that mutations at this position block splicing both *in vivo* and *in vitro* as transcripts harbouring T9 mutations fail to form spliceosomes (Vijayraghavan *et al.*, 1986). Differences at position C9 appear to be a consequence of the nature of the nucleotide substitution, with an A nucleotide resulting in the complete inhibition of splicing and a T nucleotide having little or no effect (Langford and Gallwitz, 1983; Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986; Seraphin and Rosbash, 1989a). Mutations at positions T10 and A11 produce modest increases in the levels of pre-mRNA (Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986), whereas mutations of the branchpoint A12 position inhibit splicing to a considerable degree (Langford and Gallwitz, 1983). In transcripts harbouring mutations at position A12, steps 1 and 2 of splicing can still proceed both *in vivo* and *in vitro* but with a drastic reduction in efficiency (Jaquier and Rosbash, 1986; Vijayraghavan *et al.*, 1986). All nucleotide substitutions tested at position A12 can form 2'-5' phosphodiester bonds with the G1 of the 5' splice site (Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986) and the intron lariats generated by the inefficient splicing of transcripts containing these mutations are extremely stable *in vivo*; probably as a consequence of poor recognition of the branchpoint by a "debranching" activity present in the cells (Jaquier and Rosbash, 1986). Because A12 mutations can still undergo a low level of step 2 of splicing whereas G1 mutations can only accomplish step 1, it seems likely that the G1 nucleotide is more important for step 2 of the reaction. Mutations at position C13 affect step 1 of splicing though not to a great extent (Fouser and Friesen, 1986). From all these observations it would appear that mutations in the UACUAAC

sequence primarily affect step 1 of splicing though mutations at position A12 can also inhibit step 2 (Fouser and Friesen, 1986). Actin gene transcripts harbouring mutations at the normal UACUAAC sequence can activate a cryptic UACUAAG sequence located just upstream of the normal site of branch formation (Cellini *et al.*, 1986b; Vijayraghavan *et al.*, 1986). The use of this cryptic site is very inefficient and indeed other studies have failed to detect its use (Fouser and Friesen, 1986).

Mutations at positions A14 and G15 (see Fig. 1.1) produce moderate decreases in the efficiency of step 1 of the splicing reaction, however they drastically inhibit step 2 *in vivo* (Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986). Production of mRNA molecules is not completely inhibited and 3' cleavage takes place at the mutated site. Consistent with these observations is the finding that transcripts containing mutations at position G15 can undergo splicing *in vitro* albeit at a reduced efficiency and can also form the 40S spliceosome complex as defined by glycerol gradient centrifugation (Vijayraghavan *et al.*, 1986; See Section 1.5). Transcripts containing only 29 nucleotides downstream of the branchpoint and missing an AG at the 3' cleavage site can still undergo step 1 of the splicing reaction (Rymond and Rosbash, 1985). These observations suggest that the AG dinucleotide is primarily required for step 2 of splicing. It has been found that the first AG downstream of the branchpoint is utilised as the 3' cleavage site, however inefficient use of other downstream AG dinucleotides has also been detected (Langford and Gallwitz, 1983; Pikielny *et al.*, 1983). Sequence comparisons of yeast introns have indicated that a pyrimidine nucleotide is preferred at the position preceding A14 though its importance in splicing is unknown (Langford and Gallwitz, 1983; Teem *et al.*, 1984).

Although large regions of certain introns can be deleted without affecting the efficiency of splicing (Gallwitz, 1982; Langford and Gallwitz, 1983; Pikielny *et al.*, 1983) other non-conserved intron sequences have been implicated in the splicing process. Mutational analysis has indicated that sequences close to the 5' splice and UACUAAC region influence the efficiency of splicing (Pikielny and Rosbash, 1985). In the *CYH2* gene transcript a sequence near the 5' end of the intron appears to exhibit a negative effect on splicing and indeed interacts

(in an undefined manner) with a complementary sequence just upstream of the branchpoint (Newman, 1987; see also Swida *et al.*, 1986).

Completion of both steps of splicing requires an exon 1 length of at least 12 nucleotides though a transcript with just 1 nucleotide in exon 1 can undergo step 1 of splicing (albeit inefficiently) (Duchene *et al.*, 1988). In the same study it has been observed that transcripts starting at the G1 of the intron cannot be spliced even though they can form a 40S spliceosome complex. A transcript containing only 10 nucleotides in exon 2 undergoes both steps of splicing (Rymond and Rosbash, 1985). Thus exon sequences are required for splicing though it may be that they simply act as a support for splicing-complex assembly.

Most yeast introns are found to be in the size range of 300-500 nucleotides, the largest being the intron of *L17a* (513 nts), however the smallest are the two introns of the *MATa1* gene (54 and 52 nts) (Teem *et al.*, 1984; Leer *et al.*, 1984; Miller, 1984). Increasing the size of introns above 551 nucleotides (Klinz and Gallwitz, 1985) or above 700-800 nucleotides (Swida *et al.*, 1986) reduces considerably the efficiency of splicing. This restriction on the size of introns may represent the spatial requirements of the yeast splicing machinery and therefore may explain the inefficient splicing of *MATa1* transcripts (Miller, 1984). The further an intron is located from the 5' cap structure the less efficiently it is spliced from its transcript (Klinz and Gallwitz, 1985). Indeed most yeast introns are located at the 5' end of their respective transcripts, usually among the first few nucleotides of the coding sequence (Leer *et al.*, 1984), though in at least one case the intron is placed in the 5' untranslated region (Mitra and Warner, 1984). This suggests that the introns are maintained and/or tolerated at the 5' end of their transcripts to enable their efficient removal.

The conserved sequences are subject to certain constraints regarding their spatial arrangements within introns. The distance separating the 5' splice site and the branchpoint is of critical importance in the splicing reaction. Distances below 44 nucleotides reduce the efficiency of splicing quite considerably (Duchene *et al.*, 1988; Kohrer and Domdey, 1988). The 3' splice site is usually situated 10-60 nucleotides downstream of the branchpoint (Teem *et al.*,

1984; Leer *et al.*, 1984) and the efficiency of splicing is reduced considerably when this distance is increased above 66 nucleotides (Cellini *et al.*, 1986a). Transcripts with as little as 8 nucleotides separating the branchpoint from the 3' splice site can still undergo steps 1 and 2 of the splicing reaction (Langford and Gallwitz, 1984).

#### 1.4.3 Comparison of Mammalian and Yeast Sequence Requirements

There appears to be a basic difference in the recognition of the branchpoint region in mammalian and yeast introns. Mammalian introns require the presence of a (Py)<sub>n</sub> tract immediately downstream of the branchpoint region to specify the branchpoint itself (Reed, 1989; Smith *et al.*, 1989) whereas in yeast introns the branchpoint is dictated by the presence of the invariant UACUAAC sequence (see Vijayraghavan *et al.*, 1986). The fact that mutations in the mammalian branchpoint region affect the efficiency of splicing suggests that this region is recognised in some way and is indeed analogous to the UACUAAC sequence in yeast (also see Section 1.5). The weak conservation of sequences at the mammalian branchpoint region (and indeed the 5' splice site) probably reflects the ability of mammalian transcripts to be alternatively spliced. The strict sequence requirement in yeast for the UACUAAC region is reflected in the fact that mammalian introns are not spliced in yeast cells (Beggs *et al.*, 1980). Conversely the yeast *RP51A* transcript is spliced *in vitro* in HeLa nuclear extracts though a cryptic UACUAAC sequence is utilised (Ruskin *et al.*, 1986). This cryptic branchpoint region is situated immediately upstream of a (Py)<sub>n</sub> tract and suggests a reason for its use in preference to the normal yeast UACUAAC sequence (see Reed, 1989). It would appear that in both species the 3' splice site is determined by the branchpoint i.e. the cleavage site is the first AG dinucleotide downstream of the branchpoint. Thus the fundamental processes of pre-mRNA splicing appear to be the same in mammals and yeast although the specification of the branchpoint differs with respect to the sequence requirements.

## 1.5 MAMMALIAN SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES

### 1.5.1 Introduction

Small nuclear ribonucleoprotein particles are a class of RNA-containing particles (snRNP) present in the nucleus of eukaryotic cells (reviewed in Busch *et al.*, 1982; Steitz, 1988). Each snRNP particle contains a specific small nuclear RNA (snRNA) which is complexed with at least 7 polypeptides common to all the major snRNPs (Bringmann and Luhrmann, 1986). In addition individual snRNPs can contain proteins which are specific to them. The major snRNPs are named after their snRNA component, the major species being U1, U2, U4, U5 and U6. The snRNAs are metabolically stable and range in size from 108-189 nucleotides. They are abundant molecules ranging from  $10^5$  to  $10^6$  copies/cell and with the exception of U6 possess a trimethylguanosine cap structure ( $m^3G$ ) at their 5' ends.

The study of the structure and function of these particles has been facilitated by the availability of certain antibodies which recognise their components. Sera taken from human patients suffering from autoimmune diseases such as systemic lupus erythematosus can contain antibodies that recognise all the snRNP particles (anti-Sm), U1 only (anti-(U1)-RNP), U2 only (anti-(U2)-RNP) or U1 and U2 (anti-(U1, U2)-RNP) (Lerner and Steitz, 1979; Lerner *et al.*, 1980; Lerner *et al.*, 1981). In addition antibodies recognising the trimethylguanosine moiety of the cap structure can precipitate all the major snRNA species (Bringmann *et al.*, 1983; Smith and Eliceiri, 1983). The U4 and U6 snRNAs are normally base-paired together and therefore exist in a single particle (Bringmann *et al.*, 1984; Hashimoto and Steitz, 1984). As a consequence of this interaction U6 (which does not contain a trimethylguanosine cap structure) can be co-immunoprecipitated using antibodies that recognise the trimethylguanosine cap. SnRNA structure and snRNP protein content appear to be highly conserved in evolutionary terms, as human anti-Sm and anti-(U1)-RNP antibodies can precipitate the snRNP particles from a number of organisms including yeast (Tollervey and Mattaj, 1987; Tollervey, 1987; Guthrie and Patterson, 1988).

### 1.5.2 Evidence for the Involvement of snRNPs in Splicing

The first proposals that the U1 snRNP was involved in splicing were based on the observation that the 5' end of the U1 snRNA contained regions complementary to the 5' and 3' splice sites of introns. It was suggested that the U1 snRNP aligned the splice sites such that the introns were excised accurately (Lerner *et al.*, 1980; Rogers and Wall, 1980). Evidence has since accumulated indicating the involvement of the U1 snRNP in the splicing process particularly its interaction at the 5' splice site.

Splicing of pre-mRNA transcripts in whole cell extracts is inhibited by both anti-Sm and anti-(U1)-RNP antibodies (Padgett *et al.*, 1983b). These two anti-sera also inhibit splicing *in vivo* when assayed in *Xenopus laevis* oocytes (Bozzoni *et al.*, 1984; Fradin *et al.*, 1984). Addition of oligodeoxynucleotides complementary to specific regions of U snRNAs to splicing extracts, can lead to the selective cleavage of the partial duplex created: this is caused by a RNase H activity endogenous to the splicing extracts and will subsequently be referred to as RNase H-directed cleavage. Experiments utilising this phenomenon are able to selectively degrade U1 snRNA and show its requirement for splicing (Kramer *et al.*, 1984; Black *et al.*, 1985; Black and Steitz, 1986; Berget and Robberson, 1986). Indeed, selective removal of the first 8 nucleotides of the U1 snRNA (the region thought to base-pair with the 5' splice site) completely abolishes splicing and thus implicates the U1 snRNA in a base-pairing interaction with the 5' splice site (Kramer *et al.*, 1984). Analysis of protected RNA fragments after immunoprecipitation with anti-(U1)-RNP antibodies and RNase T1 digestion shows that the U1 snRNP protects the 5' splice site (Mount *et al.*, 1983; Chabot and Steitz, 1987b). Direct genetic evidence for this base-pairing interaction was obtained when it was shown that 5' splice site mutations which inhibit splicing can be suppressed by complementary mutations in the U1 snRNA (Zhaung and Weiner, 1986). However because another mutation at a different position in the 5' splice site cannot be suppressed in a similar manner it has been suggested that recognition of the 5' splice site is dependent on other contributing factors. This analysis correlates well with similar experiments carried out in yeast (Seraphin and Rosbash, 1989a). In

this respect it has been observed that the binding of the U1 snRNP to the 5' splice site requires other factors (Tatei *et al.*, 1987; Zapp and Berget, 1989). Recently it has been shown that the C protein (which is specific to the U1 snRNP) is required for the binding of the U1 snRNP to the 5' splice site (Heinrichs *et al.*, 1990). The U1 snRNP was shown to bind to cryptic 5' splice sites even though they were not necessarily utilised as sites of cleavage (Nelson and Green, 1988; Yuo and Weiner, 1989b). It seems likely that 5' splice site recognition and cleavage have different requirements though the U1 snRNA appears to play a role in both. Analysis of 5' splice sites from a variety of introns of evolutionary diverse organisms shows that although the geometry of interaction is variable, 5' cleavage of the pre-mRNA always occurs opposite the junction between residues 9 and 10 of the U1 snRNA (Jacob and Gallinaro, 1989). This suggests that 5' cleavage is determined by the interaction with the U1 snRNP.

Further evidence for the involvement of the U1 snRNP in splicing comes from the analysis of splicing complexes. It has been shown that the U1 snRNP is present in the large ribonucleoprotein complex (spliceosome) in which splicing takes place and it is also present in other complexes related to splicing (see Section 1.5.3) (Grabowski *et al.*, 1985; Frendewey and Keller, 1985; Bindereif and Green, 1986, 1987; Chabot and Steitz, 1987a; Zillman *et al.*, 1987, 1988; Reed *et al.*, 1988; Agris *et al.*, 1989; Tazi *et al.*, 1989; Zapp and Berget, 1989). Indeed in complementation experiments using extracts depleted in specific snRNPs, the U1 snRNP was required for step 1 of the splicing reaction (Winkelman *et al.*, 1989).

Similar experiments to those described above have yielded results indicating the involvement of the U2 snRNP in splicing. RNase H-directed cleavage of U2 snRNA inhibits splicing *in vitro* (Black *et al.*, 1985; Krainer and Maniatis, 1985; Berget and Robberson, 1986; Black and Steitz, 1986). Injection of oligonucleotides complementary to the U2' snRNA into the nuclei of *Xenopus* oocytes also inhibit splicing thus providing evidence for the U2 snRNP function in splicing *in vivo* (Pan *et al.*, 1989). Immunoprecipitation with anti-(U2)-RNP antibodies followed by RNase T1 digestion reveals an interaction of the U2 snRNP with the branchpoint region (Black *et al.*, 1985). This interaction is dependent on the presence of the



(Py)<sub>n</sub> tract and to a lesser extent the AG dinucleotide present at the 3' splice site region of mammalian introns (Chabot and Steitz, 1987a). Hybridisation of biotinylated 2'-OMe RNA oligonucleotides to a region of the U2 snRNA postulated to base-pair with the branchpoint region completely inhibits splicing *in vitro* by preventing the binding of the U2 snRNP to the pre-mRNA (Lamond *et al.*, 1989). This supports the hypothesis that the U2 snRNA base-pairs with the branchpoint region. Direct genetic evidence for such an interaction has been obtained by showing shown that mutations in the branchpoint region can be suppressed by complementary mutations in the U2 snRNA (Wu and Manley, 1989) and indeed that cryptic branchpoints can be recognised more efficiently when the U2 snRNA is altered to exhibit a greater degree of complementarity to the cryptic site (Zhuang and Weiner, 1989). The binding of the U2 snRNP to the branchpoint region is dependent on the binding of another factor, U2 associated factor (U2AF), to the (Py)<sub>n</sub> tract (Ruskin *et al.*, 1988). However binding of the U2 snRNP to the branchpoint exhibits a degree of sequence preference (Nelson and Green, 1989) and indeed splicing is more efficient when the branchpoint region is complementary to the U2 snRNA branchpoint-binding sequence (Zhuang *et al.*, 1989). It has been suggested that in contrast to yeast, the base-pairing interaction between the mammalian U2 snRNP and the branchpoint region may be an optional interaction as some branchpoints exhibit no complementarity to the U2 snRNA branchpoint-binding sequence at all (Zhuang and Weiner, 1989). Different domains of the U2 snRNA appear to be required at different stages in splicing. Hybridisation of biotinylated 2'-OMe RNA oligonucleotides to the 5' end of the U2 snRNA allows the U2 snRNP to bind to the pre-mRNA but inhibits formation of an active spliceosome (Barabino *et al.*, 1989; Lamond *et al.*, 1989).

The U2 snRNP has also been identified as a component of a pre-splicing complex (Grabowski and Sharp, 1986; Konarska and Sharp, 1986; Lamond *et al.*, 1987), the spliceosome (references just indicated; Chabot and Steitz, 1987a; Bindereif and Green, 1987; Lamond *et al.*, 1988; Reed *et al.*, 1988; Agris *et al.*, 1989; Zapp and Berget, 1989) and a post-splicing complex (Barabino *et al.*, 1989). In complementation experiments utilising extracts devoid of U snRNPs,



the U2 snRNP is required for step 1 of splicing (Winkermann *et al.*, 1989).

RNase H-directed cleavage of the U4 and U6 snRNAs inhibits splicing *in vitro* (Berget and Robberson, 1986; Black and Steitz, 1986). The use of 2'-OMe RNA oligonucleotides which hybridise to specific regions of the U4 and U6 snRNAs also inhibits splicing *in vitro* (Blencowe *et al.*, 1989). In complementation experiments a reconstituted U4/U6 snRNP particle is required for splicing *in vitro* (Pikielny *et al.*, 1989) and Winkermann *et al.*, (1989) have demonstrated the requirement for the U4/U6 snRNP in the first step of splicing. Further evidence for the involvement of the U4/U6 snRNP in splicing comes from the observation that it is present in splicing complexes (Konarska and Sharp, 1986, 1987; Bindereif and Green, 1987; Lamond *et al.*, 1987, 1988; Reed *et al.*, 1988; Zillman *et al.*, 1988).

In RNase protection experiments the 3' splice site is protected from digestion by a factor that is precipitable with anti-Sm and anti-m<sup>3</sup>G antibodies, and that is also resistant to digestion with micrococcal nuclease (Chabot *et al.*, 1985). The U5 snRNP is the only major snRNP that is resistant to micrococcal nuclease leading to the suggestion that it is associated with the 3' splice site. This proposal is further supported by the observation that a protein which binds directly to the (Py)<sub>n</sub> tract, the intron-binding protein (IBP), co-purified with the U5 snRNP (Tazi *et al.*, 1986; see also Gerke and Steitz, 1986). Recently it has been shown that the U5 snRNP does interact with the 3' splice site region and requires the IBP protein to do so (Tazi *et al.*, 1989). Further evidence for the involvement of the U5 snRNP in splicing is provided by the observation that it is present in splicing complexes (Konarska and Sharp, 1985; Grabowski and Sharp, 1986; Bindereif and Green, 1987; Lamond *et al.*, 1987, 1988; Reed *et al.*, 1988; Agris *et al.*, 1989; Zapp and Berget, 1989). In complementation studies performed *in vitro* the U5 snRNP is required for efficient completion of both steps in splicing, however it appears to have a greater influence on step 2 (Winkermann *et al.*, 1989).

The snRNP particles are the focus of intense biochemical and genetic analyses. The protein constituents of the snRNPs have been determined (Luhrmann, 1988). Each particle contains a group of 7 proteins that is common to all the major snRNPs, these are called B, B', D, D', E, F and G (Bringmann and Luhrmann, 1986). Due to the fact that human antibodies specific for these particles can immunoprecipitate the analogous snRNP particles from a wide variety of organisms, it is likely that the relevant determinants (i.e. proteins, RNA and/or protein/RNA structures) are highly conserved in evolutionary diverse species (Tollervey and Mattaj, 1987; Palfi *et al.*, 1989). The transcripts encoding the B and B' proteins appear to be generated by alternative splicing of the same pre-mRNA molecule (van Dam *et al.*, 1989) and the E protein has regions of amino acid homology to a yeast mitochondrial ribosomal protein (Stanford *et al.*, 1988).

The snRNA molecules U1, U2, U4, and U5 contain the so-called "domain A" motif which consists of the sequence Pu-A-(U)<sub>n</sub>-G-Pu (n>3) situated in a single-stranded region flanked by double-stranded regions (Busch, 1982). Extensive nuclease digestion of snRNP particles results in the protection of the domain A site by a core group of proteins namely D, E, F and G (Liautard *et al.*, 1982). These core proteins assemble in the absence of the snRNA (Fisher *et al.*, 1985) suggesting that it is this core assembly that recognises the domain A motif. This motif is required for the transport of newly assembled cytoplasmic snRNP particles to the nucleus (Mattaj and DeRobertis, 1985) and can confer Sm-precipitability on RNA molecules previously unprecipitable, suggesting that the proteins and/or RNA/protein structure at this region is the Sm-determinant (Mattaj, 1986).

The U1 snRNP has 3 unique protein constituents (U1-A, 70K and C), U2 has 2 (A' and B"), a portion of the U5 snRNPs have at least 7 unique proteins associated with them (Bach *et al.*, 1989) one of which is the human homologue of the yeast PRP8 protein (Anderson *et al.*, 1989). The U6 snRNP may also have a unique protein of MW 50kD associated with it at certain developmental stages in *Xenopus laevis* (Hamm and Mattaj, 1989b). The U1-A and

U2-B" proteins share extensive regions of homology (Sillekens *et al.*, 1987) and their modes of interaction with their respective snRNAs appear to be similar suggesting that they perform similar functions (Lutz-Reyeremuth and Keene, 1989). The proteins unique to the U1 snRNP appear to be less tightly associated with the particle than the common proteins (Bach *et al.*, 1990). However the U1-70K and U1-A proteins contain a highly conserved RNA-binding domain (see Mattaj, 1989) and when assayed *in vitro* contact the U1 snRNA at stem loops I and II respectively (Scerly *et al.*, 1989; Surowy *et al.*, 1989; Query *et al.*, 1989). The U1-C protein does not contain the RNA-binding consensus (Sillekens *et al.*, 1988) and does not appear to contact the RNA (Bach *et al.*, 1990), yet it may potentiate the base-pairing interaction of the U1 snRNP with the 5' splice site (Heinrichs *et al.*, 1990).

U1 snRNA molecules containing mutations in stem loops I, II, III and domain A inhibit splicing *in vivo* when introduced into the nuclei of *Xenopus* oocytes (Hamm *et al.*, 1990; see also Yuo and Weiner, 1989a). This suggests that specific proteins normally associated with these regions are required for splicing and/or snRNP assembly and transport. However it could also indicate a U1 snRNA sequence requirement for splicing, irrespective of the associated factors.

In contrast to the mutations indicated above, U2 snRNA molecules harbouring deletions of the regions that bind the A' and B" proteins exhibit no effect on splicing when assayed in *Xenopus* oocytes (Hamm *et al.*, 1989). This suggests that these proteins have little or no effect on splicing *in vivo*.

Very little is known about the function of the other snRNP proteins though the B and B' proteins possess an endonucleolytic activity which is inactive in native snRNP particles (Temsamani *et al.*, 1989).

The elucidation of the structure/function relationships between snRNP proteins and their respective snRNAs has been greatly facilitated by the development of *in vitro* reconstitution techniques for the snRNP particles. Initially this has been developed with the U1 snRNP (Hamm *et al.*, 1987; Patton *et al.*, 1987; see also Patton *et al.*, 1989) but has recently been used

to identify the regions of U6 required for U4 interaction and spliceosome assembly (Bindereif *et al.*, 1990). Indeed reconstitution has been attempted for all the major snRNPs and an active U4/U6 particle has been formed (Pikielny *et al.*, 1989).

## 1.6 OTHER FACTORS INVOLVED IN MAMMALIAN SPLICING

In complementation studies using biochemically fractionated extracts, non-snRNP components required for splicing *in vitro* have been identified (Furneau *et al.*, 1985; Krainer, 1988). Factors termed SF4B and SF4A are required for steps 1 and 2 of splicing respectively (Krainer and Maniatis, 1985). In other experiments 6 fractions required for splicing have been identified (Kramer *et al.*, 1987). Indeed from these studies fractions SF1 and SF3 along with the U2 snRNP are required for the formation of a pre-splicing complex (Kramer, 1988). Two non-snRNP components which co-purify with *de novo* spliceosomes have also been identified (Conway *et al.*, 1989). However, due to the different experimental techniques utilised by individual investigators, the relationship of the components required for splicing remains unclear. Conditions have recently been defined in which fully operable *in vitro*-assembled spliceosomes have been formed and this should accelerate the dissection of the functional relationships of the constituents of the spliceosome (Parent *et al.*, 1989).

A monoclonal antibody raised against hnRNP C protein inhibits step 1 of splicing *in vitro*, however it does not prevent the assembly of the 60S spliceosome complex (see Section 1.7) (Choi *et al.*, 1986). Immunodepletion of the hnRNP C-type protein from the splicing extracts results in the complete inhibition of spliceosome formation. The hnRNP A, C and D proteins appear to bind to the 3' splice site region of introns (Swanson and Dreyfus, 1988). These observations suggest the involvement of the hnRNP C, A and D proteins in pre-mRNA splicing.

A factor termed the U2 auxillary factor (U2AF) is required for the binding of the U2 snRNP to the branchpoint region (Ruskin *et al.*, 1988). U2AF binds to the 3' splice site region

and may be present in the SF2 fraction previously identified by Kramer (Kramer *et al.*, 1987).

Two independent studies have identified a protein that binds to the (Py)<sub>n</sub> tract located upstream of the 3' splice site (Tazi *et al.*, 1986; Gerke and Steitz, 1986). It is likely that the differences observed in the MW (100 and 70kD respectively) of the proteins isolated reflect the different isolation procedures used, and thus both studies have probably identified the same protein. This intron-binding protein (IBP) co-purifies with the U5 snRNP (Tazi *et al.*, 1986) and is immunoprecipitable with anti-Sm antibodies (Gerke and Steitz, 1986). IBP appears to bind to the 3' splice site in the absence of ATP and this may constitute one of the earliest events in the assembly of the spliceosome (Tazi *et al.*, 1989).

Recently a protein, present only in the oocyte stage of *Xenopus* development, has been identified (Hamm and Mattaj, 1989). This 50kD protein appears to associate with U6 snRNPs that are not complexed with U4 snRNPs and may be involved in the transport of the U6 snRNP to the nucleus rather than in splicing. Factors other than the U1 snRNP have been implicated in the recognition of the 5' splice site (Tatei *et al.*, 1987; Zapp and Berget, 1989).

## 1.7 MAMMALIAN SPLICING COMPLEXES

### 1.7.1 Isolation and Identification

To align the 5' and 3' splice sites and account for the involvement of the major snRNP particles it was postulated that splicing took place in a large ribonucleoprotein complex. This was first identified as 60S and 50S complexes in glycerol and sucrose gradient sedimentation analysis respectively (Grabowski *et al.*, 1985; Frendewey and Keller, 1985; Bindereif and Green, 1986). These complexes contain the intermediates of the splicing reaction, thereby defining them as active so-called spliceosomes. Formation of the active spliceosome proceeds in a stepwise manner requiring all the major U snRNPs and ATP (Grabowski *et al.*, 1985; Frendewey and Keller, 1985; Grabowski and Sharp, 1986; Bindereif and Green, 1986;

Zillman *et al.*, 1987). The 5' and 3' splice site consensus sequences are also necessary for the formation of the spliceosome, although transcripts containing deletions or mutations at either the 5' or 3' splice site can still form smaller complexes (Grabowski *et al.*, 1985; Frendewey and Keller, 1985). Pre-spliceosome complexes (35-40S) containing snRNP particles have also been identified and appear to form in an ATP-independent manner requiring the presence of the 3' splice site consensus sequence (Grabowski *et al.*, 1985; Frendewey and Keller, 1985; Bindereif and Green, 1986).

Splicing complexes have been further analysed utilising polyacrylamide gel electrophoresis techniques (PAGE). Initially two splicing complexes, arising in a sequential manner, were defined (Konarska and Sharp, 1986, 1987). The so-called complex A requires the 3' splice site region for its formation and gives rise to complex B which requires the 5' splice site. In further studies using a different pre-mRNA transcript as the substrate, 3 complexes were resolved which also arise in a sequential manner. The so-called complex  $\alpha$  gives rise to  $\beta$  which precedes complex  $\gamma$  (Lamond *et al.*, 1987). Complexes B and  $\gamma$  contain the splicing intermediates and are therefore defined as spliceosomes. Complex B can be resolved into two components ( $\beta$  and  $\gamma$ ) by this analysis, one of which (complex  $\gamma$ ) lacks the U4 snRNP. All the complexes contain the U2 snRNP and require ATP for their formation. The U4/U6/U5 complex (see 1.7.1) joins the forming spliceosome in the transition from complex A to B.

#### 1.7.2 Temporal Order of snRNP Addition in Spliceosome Assembly

Having observed that active spliceosome formation arises in a stepwise manner, it became important to define the stages at which the splicing factors (i.e. snRNPs and other components) join the spliceosome. There has been some controversy about the earliest events in spliceosome assembly. The U1 snRNP is detected in early splicing complexes (Zillman *et al.*, 1987, 1988) and binds to the 5' splice site, in an ATP-independent manner, prior to the binding of the U2 snRNP (Bindereif and Green, 1987; Tazi *et al.*, 1989). However the U1 snRNP remains undetected in a number of splicing-complex studies (Grabowski and Sharp, 1986; Konarska and Sharp, 1987; Lamond *et al.*, 1987, 1988). These anomalous results are probably

a consequence of the different experimental procedures used. The IBP protein (hence implicating the U5 snRNP, see Section 1.6) binds to the (Py)<sub>n</sub> tract, located at the 3' splice site region, at an early stage in spliceosomal assembly and in the absence of ATP (Tazi *et al.*, 1989). It has been suggested that the U1 snRNP and the IBP protein (U5 snRNP) recognise the 5' and (Py)<sub>n</sub> tract respectively and that this constitutes the earliest event in spliceosome assembly (Tazi *et al.*, 1989). Consistent with this hypothesis is the observation that the 5' and 3' splice sites collaborate in early complex formation (Lamond *et al.*, 1987) and that the U1 snRNP appears to interact with the 3' splice site at early stages of spliceosome formation (Zillman *et al.*, 1987, 1988; see also Tatei *et al.*, 1987).

Pre-splicing complexes containing the U2 snRNP and requiring ATP for their formation have been identified (Grabowski and Sharp, 1986; Konarska and Sharp, 1986, 1987; Lamond *et al.*, 1987, 1988); this suggests that the binding of the U2 snRNP is the next step in spliceosome assembly. It has been postulated that the binding of the U1 snRNP to the 5' splice site followed by the interaction of the U2 snRNP at the branchpoint is the preferred order of events (Bindereif and Green, 1987). Consistent with this suggestion is the observation that complexes can form on a transcript lacking either the 5' or 3' splice site consensus sequences (Grabowski *et al.*, 1985; Frendewey and Keller, 1985). However in some cases it has been observed that formation of a stable complex at the 5' splice site requires the presence of the 3' splice site sequences (Ruskin and Green, 1985), the U1 snRNP, the U2 snRNP and ATP (Zapp and Berget, 1989). These pre-splicing complexes i.e. those forming prior to the addition of the U4/U6/U5 complex, probably correspond to complex A as defined by PAGE.

The U5 snRNP interacts, in an ATP-dependent manner, with the U4/U6 snRNP particle (Pinto and Black, 1989). This complex then joins the pre-splicing complexes described above (Konarska and Sharp, 1987; Lamond *et al.*, 1988). It has been previously observed, by PAGE analysis, that the U4 snRNP leaves the spliceosome prior to step 1 of the splicing reaction (Lamond *et al.*, 1988; Pikielny *et al.*, 1989). However recent data, using biotinylated 2'-OMe RNA oligonucleotides complementary to the U4 snRNA, suggests that although a



conformational change is likely to take place, the U4 snRNP remains associated with the spliceosome (Blencowe *et al.*, 1989). The addition of the U4/U5/U6 complex to the U1/U2 pre-splicing complex probably defines the active spliceosome.

It has been proposed that the U5 snRNP acts at two stages in splicing i.e. at the initial recognition of the 3' splice site sequences and at a later stage when it is complexed with the U4/U6 snRNP particle (Tazi *et al.*, 1989). The U5 snRNP acting at the initial stage may be replaced by other 3' splice site binding factors such as U2AF and become available for the formation of the U4/U5/U6 complex. This hypothesis is consistent with the observations of Winkelmann *et al.*, (1989).

Post-splicing complexes containing the excised intron and the U2, U5 and U6 snRNPs have been observed (Konarska and Sharp, 1987; Zillman *et al.*, 1988).

## 1.8 YEAST SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES

### 1.8.1 Introduction

Yeast cells contain at least 23 species of snRNAs which exhibit a varied degree of abundance, size and complexity (Riedal *et al.*, 1986; see also Tollervey, 1987). Initial experiments indicated that a large proportion of these snRNA molecules are not required for cell viability (reviewed in Guthrie, 1986a, 1986b; Patterson and Guthrie, 1988), indeed a haploid strain in which five snRNA genes are deleted exhibits growth properties similar to wild-type cells (Parker *et al.*, 1988). Certain yeast snRNAs become immunoprecipitable with anti-Sm antibodies when introduced into *Xenopus* oocytes, suggesting the evolutionary conservation of the factors that recognise the Sm-binding site (Riedal *et al.*, 1987). The yeast homologues of the major mammalian U1, U2, U4, U5 and U6 snRNA molecules have been identified and their genes cloned (Patterson and Guthrie, 1988). Similar to their mammalian homologues the major yeast snRNAs possess a trimethylguanosine cap, contain the highly conserved domain A motif required for the binding of the Sm determinant (see Section 1.5) and share sequence and structural



homologies to their mammalian counterparts. The U6 snRNA of yeast, as with its mammalian homologue, has a modified cap structure (distinct from the trimethylguanosine cap) and does not possess the Sm binding site (domain A). The major yeast snRNAs are essential for viability.

In contrast to their mammalian counterparts the major yeast snRNA genes are present in single-copy in the genome and the snRNA molecules are found in relatively low abundance in the cell (200-500 copies/cell). In addition considerable size variation can exist between the yeast snRNAs and their mammalian counterparts.

### 1.8.2 Evidence for the Involvement of Yeast snRNPs in Splicing

Micrococcal nuclease treatment of yeast splicing extracts inhibits splicing *in vitro* (Cheng and Abelson, 1986). This suggested that snRNP particles participate in splicing in yeast. Anti-m<sup>3</sup>G antibodies which recognise the 5' cap structure of the snRNAs (except U6) are able to immunoprecipitate the splicing intermediates suggesting the presence of snRNP particles in the active spliceosome (Cheng and Abelson, 1986). Indeed splicing complexes resolved by native gel electrophoresis contain the U2, U4, U5 and U6 snRNAs (Pikielny and Rosbash, 1986; Pikielny *et al.*, 1986; Cheng and Abelson, 1987).

The yeast U1 snRNA, at 569 nucleotides, is more than 3 times the size of its mammalian counterpart. The 5' ends of the yeast and mammalian U1 snRNA are highly conserved and it is this region that was postulated to base-pair with the 5' splice site (Kretzner *et al.*, 1987). RNase H-directed cleavage of the U1 snRNA (Legrain *et al.*, 1988) and in particular the 5' end of the molecule (Kretzner *et al.*, 1987) results in the complete inhibition of splicing *in vitro*. The latter observation is consistent with the postulated base-pairing interaction between U1 snRNA and the 5' splice site. *In vivo* analysis indicated that mutations of the 5' splice site, which inhibit splicing, are suppressed by the complementary mutation in the U1 snRNA, thus proving the base-pairing interaction (Seraphin *et al.*, 1988; Siliciano and Guthrie, 1988; Seraphin and Rosbash, 1989a). Further evidence for the participation of the U1 snRNP in splicing comes from the observation that it is present in splicing complexes (Kretzner

*et al.*, 1987; Ruby and Abelson, 1988; Seraphin and Rosbash, 1989b).

The yeast U2 snRNA, at >1000 nucleotides, is ~6 times the length of its mammalian counterpart (Ares, 1986). RNase H-directed cleavage of U2 snRNA results in the inhibition of splicing *in vitro* and suggests the involvement of the U2 snRNA in splicing (Legrain *et al.*, 1988; McPheeters *et al.*, 1989). Strikingly, large internal regions of the U2 snRNA are inessential for splicing *in vivo* (Iguel and Ares, 1988) and indeed the human and rat U2 snRNAs can functionally replace the yeast U2 snRNA in yeast extracts (McPheeters *et al.*, 1989). In snRNP reconstitution studies it has been shown that only 123 nucleotides present at the 5' end of the yeast U2 snRNA molecule are required for splicing *in vitro* (McPheeters *et al.*, 1989). Mutations of the UACUAAC sequence which inhibit splicing are suppressed by complementary mutations in the U2 snRNA both *in vivo* (Parker *et al.*, 1987) and *in vitro* (McPheeters *et al.*, 1989), indicating a base-pairing interaction between the U2 snRNA and the branchpoint. The U2 snRNA has been found in splicing complexes, further indicating the participation of the U2 snRNP in splicing in yeast (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Ruby and Abelson, 1988).

The yeast U4 and U6 snRNAs are involved in an extensive base-pairing interaction suggesting that they are components of the same snRNP particle as observed with their mammalian counterparts (Siliciano *et al.*, 1987; Brow and Guthrie, 1988). The U6 snRNA is the most highly conserved snRNA exhibiting 75% identity between yeast and humans (Brow and Guthrie, 1988). RNase H-directed cleavage of the U6 snRNA analysed *in vitro*, results in the inhibition of splicing and an accumulation of early splicing complexes (Fabrizio *et al.*, 1989). Unlike the human U2 snRNA, the human U6 snRNA cannot functionally replace the yeast U6 snRNA *in vitro*. This observation is surprising considering the high degree of sequence and structural conservation exhibited by the U6 snRNA molecules. The U4 and U6 snRNAs are constituents of splicing complexes further indicating the participation of the U4/U6 snRNP particle in splicing (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Ruby and Abelson, 1988).

Depletion of the U5 snRNA, achieved *in vivo*, leads to the accumulation of pre-mRNA

and the IVS-E2 lariat intermediate (Patterson and Guthrie, 1987). These observations suggest that the U5 snRNP is involved in splicing in yeast and prompted the proposal that it functioned at the two steps in splicing. A similar proposal has been advocated for the function of the U5 snRNP in mammalian splicing (Winkelman *et al.*, 1989; see also Tazi *et al.*, 1989). The yeast U5 snRNP is also present in splicing complexes (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Ruby and Abelson, 1988).

### 1.8.3 Yeast snRNP Proteins

Analysis of yeast snRNPs has been hindered by their low abundance. The existence of a batch of temperature-sensitive ( $t^s$ ) mutants defective in precursor RNA processing (*prp*; see Section 1.9) has facilitated the identification of 2, possibly 3, snRNP proteins. Antibodies have been raised against the protein products of the *PRP* genes and used to immunoprecipitate snRNP particles and splicing complexes from yeast extracts. Indeed antibodies raised against PRP4 and 11 inhibit splicing *in vitro* suggesting their requirement for splicing (Chang *et al.*, 1988; Banroques and Abelson, 1989), and immunodepletion of yeast extracts using anti-PRP8 antibodies abolishes splicing activity indicating its participation in pre-mRNA splicing (Jackson *et al.*, 1987). Using these sort of approaches, PRP8 has been identified as a component of the U5 snRNP (Lossky *et al.*, 1987; Jackson *et al.*, 1988) and the spliceosome (Whittaker *et al.*, 1990); and PRP4 identified as a component of the U4/U6 snRNP (Petersen-Bjorn *et al.*, 1989; Banroques and Abelson, 1989). RNase digestion of the U4/U6 snRNP followed by immunoprecipitation of the protected RNA fragments has revealed that PRP4 is primarily associated with the 5' end of the U4 snRNA, though there is also a loose association with the 3' end of the U6 snRNA (Xu *et al.*, 1990). PRP11 may not be a snRNP protein, however it is associated with the 40S spliceosome complex (see Section 1.11) and a putative 30S pre-splicing complex (Chang *et al.*, 1988). Immunofluorescence studies have placed PRP11 at the periphery of the nuclear membrane (Chang *et al.*, 1988).

## 1.9 GENETIC AND BIOCHEMICAL APPROACHES TO THE ISOLATION OF FACTORS INVOLVED IN PRE-mRNA SPLICING IN YEAST

The *prp* mutants comprise a set of temperature-sensitive ( $t^s$ ) mutants originally isolated from a screen of 400  $t^s$  mutants assayed for defects in macromolecular synthesis (Hartwell, 1967). The basis of this screen was the incorporation of  $^{13}\text{C}$ -adenine and  $^3\text{H}$ -lysine into DNA/RNA and protein respectively, when a  $t^s$  mutant was incubated at the non-permissive temperature. Mutant cultures which exhibited a protein:RNA (P/R) ratio  $>4$  after 3-6 hours incubation at the restrictive temperature were classed as being defective in RNA metabolism. Twenty-three such mutants were later placed into 9 complementation groups termed *prp2-prp11* (previously called *rna2-rna11*: N.B. *prp10* and *prp11* are the same complementation group) (Hartwell *et al.*, 1970). These mutants were subsequently shown to be defective in pre-mRNA processing (Rosbash *et al.*, 1981; Teem and Rosbash, 1983; Lustig *et al.*, 1986) and not in the rate of transcription (Kim and Warner, 1983). Northern blot analysis of RNA prepared from the *prp2-prp11* mutant strains incubated at the non-permissive temperature, shows that they all accumulate pre-mRNA (Jackson, 1987). This suggests that the *prp2-prp11* mutants are defective in pre-mRNA splicing at the non-permissive temperature. Further proof of the involvement of the *RP2-PRP11* proteins in splicing has been obtained from the observation that splicing extracts made from the *prp3*, *prp4*, *prp5*, *prp7*, *prp8* and *prp11* mutants cannot form spliceosomes at the restrictive temperature *in vitro*, suggesting the involvement of the PRP proteins in spliceosome formation (Lustig *et al.*, 1986). In contrast extracts made from *prp2* mutants can form a 40S-like spliceosome at the restrictive temperature, however this complex cannot undergo step 1 of the splicing reaction. The various mutant extracts can complement each other at the elevated temperature suggesting that specific components are inactivated in the different extracts and that these components are freely exchangeable *in vitro*.

These results suggest that PRP2 functions before step 1 of the splicing reaction. The presence of extra plasmid-borne copies of PRP3 is able to relieve the  $t^s$  phenotype of a *prp4*

mutant, suggesting that PRP3 and PRP4 perform a similar function (Last *et al.*, 1987). Null alleles of *PRP4* are not suppressed by the overexpression of PRP3 suggesting that PRP3 and PRP4 interact with each other, perhaps in the same complex (snRNP particle) (Last and Woolford, 1986). PRP6 and PRP9 may function at the commitment stage of a pre-mRNA transcript to splicing (Legrain and Rosbash, 1989).

The accumulation of pre-mRNA exhibited by *prp* mutants incubated at the non-permissive temperature has been utilised as the basis of a screen of ~1000 *t<sup>s</sup>* mutants for defects in splicing (Vijayraghavan and Abelson, 1989). RNA was prepared from these mutants, after a 2 hour incubation at the non-permissive temperature, and analysed by Northern blot analysis. Thirty-seven *prp* mutants were identified by this screen and they can be placed into 4 distinct groups based on their Northern blot phenotype. These are; (a) those that accumulate pre-mRNA only; these include new allelic mutants of the pre-existing *prp2*, *prp3*, *prp6* and *prp9* mutants, and those placed into new complementation groups, *prp19*, *prp20*, *prp21*, *prp23*, *prp24* and *prp25*, (b) those that accumulate pre-mRNA and the intron-lariat, *prp22*, (c) those that accumulate the IVS-E2 lariat intermediate only, *prp17* and *prp18*, (d) those that accumulate the intron lariat only, *prp2* and *prp27*.

Eleven new complementation groups have been identified (*prp17-prp27*). The fact that *prp25* and *prp26* are not temperature sensitive suggests that certain mutations reduce the efficiency of splicing *in vivo* rather than completely inhibiting it. In this respect it has been proposed that splicing itself is not the rate-limiting step in pre-mRNA processing (Pikielny and Rosbash, 1985) and it is therefore conceivable that splicing mutants showing the growth phenotype of *prp25* and *prp26* can exist.

Heat-inactivated extracts made from a *prp18* mutant are unable to undergo the second step of splicing and accumulate the 40S spliceosome (see Section 1.10) containing the splicing intermediates (Vijayraghavan and Abelson, 1990). Spliceosomes isolated from the extracts made from a *prp18* mutant can be complemented (for step 2) by extracts from other *prp* mutants, and with an extract treated with micrococcal nuclease. In addition, efficient complementation

requires ATP. These observations suggest respectively that PRP18 is required for step 2 of the splicing reaction, is readily exchangeable *in vitro*, may not be a snRNP component (though the U5 snRNP is resistant to micrococcal nuclease and therefore PRP18 may be a component of this snRNP) and that ATP is required for the second step of splicing. Thus PRP18 is necessary but not sufficient for step 2 of splicing.

A more direct approach to the isolation of putative splicing factors is the generation of extragenic suppressor mutants. The *SRN1* mutant suppresses the *t<sup>s</sup>* phenotype of a strain harbouring the *prp2* and *prp6* mutations (Pearson *et al.*, 1982). *SRN1* suppresses *prp1*, 2, 3, 4, 5, 6 and 8 either singly or in pairs, however its function remains unknown. The *SRN2* mutation suppresses the *prp2* mutant phenotype (Last *et al.*, 1987). Its interaction with *prp2* is suggested by the fact that *SRN2* does not suppress null alleles of *PRP2*. The *SPP81* mutant is a cold-sensitive suppressor of the *t<sup>s</sup>* phenotype of the *prp8-1* mutant and based on the predicted protein sequence, the *SPP81+* gene appears to encode a putative ATP-dependent RNA helicase (D. Jamieson, pers. comm.). The generation of conditional-lethal suppressors in an alternate cold-sensitive, heat-sensitive manner is a powerful genetic approach to the isolation of interacting factors. The conditional-lethal phenotype of the mutants makes genetic manipulation easier and facilitates the isolation of the wild-type gene of interest. The *prp16* mutant is a *trans*-acting suppressor of a branchpoint mutation (Couto *et al.*, 1987). The *PRP16* gene has been cloned and sequenced and it has been proposed that PRP16 is a putative ATP-binding protein involved in the fidelity of branchpoint recognition (Burgess *et al.*, 1990).

Biochemical fractionation of yeast splicing extracts has identified 3 fractions (I, II and III) that are required for splicing *in vitro* (Cheng and Abelson, 1986). Fractions I and II together can perform step 1 of the splicing reaction, however fraction III is necessary for step 2. At least 2 factors required for splicing which are extrinsic to the spliceosome were identified; *c<sub>n</sub>* and *b<sub>n</sub>* are required for 3' and 5' splice site cleavage respectively (Cheng and Abelson, 1987). In the presence of ATP fraction I can complement heat-inactivated *prp18* extracts (Vijayraghavan and Abelson, 1990) and indeed *c<sub>n</sub>* is present in fraction I, therefore it is possible that *c<sub>n</sub>* is

PRP18 (which is necessary for step 2). An intrinsic component of the spliceosome, cIII, is required for step 2 of splicing and this has also been identified by biochemical fractionation (Cheng and Abelson, 1987).

## 1.10 YEAST SPLICING COMPLEXES

### 1.10.1 Isolation and Identification

Splicing complexes in yeast have been distinguished by glycerol gradient centrifugation (Brody and Abelson, 1985). A 40S complex has been identified whose efficient formation requires ATP and the presence of the correct 5' and branchpoint consensus sequences. The presence of the IVS-E2 lariat and exon 1 splicing intermediates in this 40S complex defines it as the active spliceosome.

Splicing complexes have also been identified by native gel electrophoresis and shown to require ATP, the 5' splice site and branchpoint consensus sequences for their efficient formation (Pikielny and Rosbash, 1986; Pikielny *et al.*, 1986). Three complexes arising in a stepwise manner have been resolved and these are termed III-I-II (Pikielny *et al.*, 1986). The U2 snRNA is present in all three complexes, U4 snRNA in I only and U5 snRNA in I and II. Complex II contains the splicing intermediates and is thus defined as the active spliceosome. A similar approach using a different experimental procedure also identified a stepwise assembly process in spliceosome formation (Cheng and Abelson, 1987). In this study the U6 snRNA is present in complexes analogous to complexes I and II above, indicating the participation of all the major snRNPs (except U1) in splicing in yeast. However, U1 snRNA base-pairs with the 5' splice site early in splicing (Seraphin *et al.*, 1988) and indeed this interaction occurs in the absence of ATP and before the binding of the U2 snRNP to the pre-mRNA (Ruby and Abelson, 1988). So-called "commitment" complexes, which form in the absence of ATP, contain the U1 snRNA, pre-mRNA and possibly other factors (but not the U2 snRNA) have been identified by native gel electrophoresis (Seraphin and Rosbash, 1989b; see also Legrain *et al.*,



1988).

In RNase H-directed cleavage experiments the biochemical requirements for spliceosome formation appear to parallel the resistance of the 5' splice site and branchpoint to digestion, indicating that these two sites are bound by factors required for spliceosome formation (Rymond and Rosbash, 1986). Protection at either site requires the presence on the pre-mRNA transcript of the other consensus sequence and is also dependent on ATP. The pattern of protection of the 5' splice site changes as the splicing reaction proceeds, and indeed correlates with the change from complex III to I/II (Rymond and Rosbash, 1986). In addition, chemical modification studies indicate that nucleotides at the 5' splice site participate in several roles in spliceosome assembly and splicing (Rymond and Rosbash, 1988).

#### 1.10.2. Temporal Order of snRNP Interaction in Spliceosome Assembly

It is now apparent that the temporal order of events leading to spliceosome formation in yeast mirrors closely the assembly of spliceosomes in mammalian cells. The U1 snRNP binds to the pre-mRNA at the 5' splice site (possibly interacting with the 3' splice site) in an ATP-independent manner (Ruby and Abelson, 1988). This is followed by the ATP-dependent binding of the U2 snRNP at the branchpoint (other as of yet unidentified factors may bind to the pre-mRNA prior to this U2 snRNP interaction, e.g. the yeast equivalent of U2AF) (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Ruby and Abelson, 1988). The U5 and U4/U6 snRNPs appear to complex together in an ATP-dependent manner prior to their combined association with the U1/U2/ pre-mRNA complex which leads to the formation of the spliceosome (Lossky *et al.*, 1988; see also Pinto and Black, 1989). As in mammalian studies using gel electrophoresis techniques the yeast U4 snRNP appears to leave the spliceosome prior to step 1 of splicing (Pikielny *et al.*, 1986; Cheng and Abelson, 1987). However the apparent exit of the U4 snRNP may only reflect a conformational change in the U4/U6 interaction which leaves the U4 snRNP less tightly associated with the spliceosome. Due to this weakened interaction the U4 snRNP may be displaced from the spliceosome as a consequence of the gel electrophoresis conditions



employed. Post-splicing complexes containing PRP8 (hence the U5 snRNP) have been identified by Whittaker *et al.*, (1990).

It has been proposed that the U4 snRNA is in fact a negative regulator of a catalytic U6 snRNA function (Brow and Guthrie, 1989). Unwinding of the U4/U6 base-pairing interaction (analogous to the conformational change observed in native gel electrophoresis) may unmask the highly conserved U6 sequence proposed to act catalytically in the splicing process. This putative catalytic function of the U6 snRNA is based on the observation that the U6 snRNA gene of *Schizosaccharomyces pombe* contains a pre-mRNA-type intron in the highly conserved region of the U6 snRNA (Tani and Ohshima, 1989). Insertion of the intron at this site is postulated to have occurred as a consequence of a mishap during a splicing reaction. This would therefore place the highly conserved region of the U6 snRNA at a site of catalytic activity in the spliceosome (Brow and Guthrie, 1989). It is also interesting to note that the U2 snRNA which itself may reside near a catalytic centre in the spliceosome (due to its interaction with the branchpoint region) has a potential base-pairing interaction with the U6 snRNA (McPheeters *et al.*, 1989).

## 1.11 PRE-mRNA SPLICING AND THE CONTROL OF GENE EXPRESSION

### 1.11.1 Alternative Splicing

Alternative splicing of pre-mRNA transcripts has emerged as an important device for regulating gene expression. An extensive discussion of this topic is beyond the scope of this introduction and the reader is referred to the comprehensive review of Smith *et al.*, 1989a.

There are several ways in which alternative splicing can generate different transcripts. These modes of alternative splicing can be classified into 5 groups; (i) transcripts in which an intron has been retained or removed, (ii) transcripts in which alternative 5' and 3' splice sites are utilised, (iii) transcripts which either initiate at or are polyadenylated at different promoter and/or polyadenylation sites respectively (this can lead to the exclusion/inclusion of exons from

the 5' and 3' ends of the transcripts), (iv) transcripts which can contain exons that are removed in a mutually exclusive manner and (v) transcripts where individual exons can be included or excluded independently of other sequences.

There are a number of ways in which alternative splicing can affect gene expression. Different (though similar) proteins encoded by alternatively spliced transcripts can be membrane bound or secreted from the cell, e.g. the immunoglobulin protein (*Ig $\mu$ m*) proceeds from being in membrane-associated form to being in a secreted form (*Ig $\mu$ s*) as a result of alternative splicing of the 3' end exons residing in the transcript that encodes it. Alternative splicing, causing early termination of translation, can erase the function of a protein, e.g. the *Sex-lethal* protein of *Drosophila*. Protein function itself may be modulated by the inclusion or exclusion of certain exons, e.g. the tyroponin-T protein, and indeed this can also lead to the production of proteins with different functions, e.g. calcitonin and calcitonin gene-related peptide (CGRP). Alternative splicing which creates heterogeneity in the 5' and 3' untranslated regions may affect translational efficiency and/or RNA stability.

Various determinants appear to be involved in splice site selection. The interaction of these determinants is not at all clear and is the object of intense study. I will briefly out-line the determinants implicated so far without discussing the mechanisms proposed to account for their involvement. What is clear is that transcripts may have different determinants which regulate their alternative splicing patterns.

(i) Some introns contain branchpoints further upstream ( $\sim 100$ nts) than would normally be expected. It has been proposed that the long (Py)<sub>n</sub> tracts found adjacent to these branchpoints are important in branchpoint recognition and may be an important feature of transcripts that are alternatively spliced (Reed, 1989; Goux-Pelletan *et al.*, 1990). The branchpoint itself can determine 5' splice site utilisation (Noble *et al.*, 1988, 1989).

(ii) Differential recognition of introns by distinctive forms of snRNPs (some of which are known to be developmentally regulated) has been postulated as an important factor in alternative splicing (see Lund and Dahlberg, 1987; Lund *et al.*, 1987). Indeed complementarity of the 5' splice site with the U1 snRNA may increase the use of that site (Lear *et al.*, 1990). It is also

known that increased complementarity between the branchpoint region and the U2 snRNA is important for splicing efficiency (Zhuang *et al.*, 1989).

(iii) Intron size appears to be important in splice site selection (Fu and Manley, 1987; see also Reed and Maniatis, 1986).

(iv) Exon sequences can also influence the choice of splice sites (Reed and Maniatis, 1986; Tsai *et al.*, 1989; Nagoshi and Baker, 1990) and indeed it has been suggested that the "definition" of the exons can determine splice site selection (Robberson *et al.*, 1990).

(v) RNA secondary or indeed tertiary structure may influence splice site choice (Khoury *et al.*, 1979; Solnick, 1985; Eperon *et al.*, 1986; Chebli *et al.*, 1989).

(vi) Introns may contain sequences (outwith the consensus sequences) that influence the selection of splice sites (Helfman *et al.*, 1990).

(vii) *Trans*-acting factors are involved in both determining (Nagoshi and Baker, 1990) and inhibiting (Inoue *et al.*, 1990) splice site utilisation.

(viii) The "first come first served" principle has been postulated based on the observations that some introns can be removed co-transcriptionally, suggesting that it is the transcriptional order of appearance of an intron that determines its removal (Beyer and Osheim, 1988).

Probably the most exciting discovery so far is that sexual differentiation in *Drosophila melanogaster* is regulated by a cascade of alternative splicing events. The epistatic hierarchy of genes involved in this cascade is well established (reviewed in Smith *et al.*, 1989). The *Sex-lethal* gene (*Sxl*) activates *transformer* (*tra*) which in conjunction with *transformer-2* (*tra-2*) then acts to establish the female mode of expression of the *double-sex* gene (*dsx*). *Sxl*, *tra* and *dsx* are all regulated at the level of splicing. Each gene can produce two transcripts which are either male or female-specific, though for *Sxl* and *tra* the male-specific transcripts are non-functional. The female-specific transcript of *Sxl* encodes a protein which has an RNA-binding consensus (Bell *et al.*, 1988) and this binds to the male-specific 3' splice site of the *tra* transcript (Inoue *et al.*, 1990). Thus *Sxl* inhibits the utilisation of the male-specific 3' splice site and so *tra* only generates female-specific transcripts. *Sxl* is thought to regulate its own female-specific expression

in a similar manner. In contrast, the choice of the female-specific splice site in the *dsx* transcript appears to be determined in a positive manner i.e. *trans*-acting factors increase its use, rather than inhibit the use of the male-specific sites (Nagoshi and Baker, 1990).

The failure to mimic the *in vivo* pattern of splicing *in vitro* (and so identify factors involved in alternative splicing) has revealed the advantages provided by the genetic system available in *Drosophila*. It seems likely that studies utilising *Drosophila* may provide more examples of systems controlled by a cascade of alternative splicing events.

#### 1.11.2 Introns: Effects on Gene Expression

Early experiments utilising viral recombinants indicated that introns are required for the production of stable cytoplasmic RNA (Gruss *et al.*, 1979; Hamer and Leder, 1979) and the efficient expression of the gene encoding the chicken ovalbumin protein (Wickens *et al.*, 1980). However, introns are not required for the expression of other viral genes (Ghosh *et al.*, 1981; Carlock and Jones, 1981; Treisman *et al.*, 1981). Recent data from a number of studies involving a variety of genes, have indicated that introns are necessary for efficient gene expression and indeed can increase expression several hundred-fold. The genes studied include the mouse ribosomal gene *RPL32* (Chung and Perry, 1989), a  $\beta$ -globin gene analysed in SV40 recombinants (Buchman and Berg, 1988), the human proliferating cell nuclear antigen gene (*PCNA*) (Ottavio *et al.*, 1990) and the alcohol dehydrogenase-1 gene (*Adh1*) of maize cells (Callis *et al.*, 1987).

This intron-dependent effect on gene expression does not appear to be a consequence of differences in the levels or rates of transcription between intron-containing (IVS(+)) and intron-less (IVS(-)) genes (Wickens *et al.*, 1980). Results obtained from nuclear run-off experiments are consistent with this view (Huang and Gorman, 1990). In the study of Huang and Gorman, steady-state levels of total nuclear RNA transcribed from an IVS(+) and IVS(-) gene, and the rate of decay of both types of RNA generated, is similar. This suggests that the intron effect on splicing is not at the level of transcription and/or RNA stability. Although the

steady-state levels of nuclear IVS(+) and IVS(-) RNA are similar, polyadenylated nuclear RNA consists of (in this study) seven times more IVS(+) RNA. This distribution of polyadenylated RNA is mirrored in the cytoplasm i.e. there is 7-fold more mRNA (poly. A+) derived from the IVS(+) gene in the cytoplasm than mRNA (poly A+) derived from the IVS(-) gene. This suggests that the rates of transport of the two RNA species are similar. Thus it has been postulated that there is a selective transition of IVS(+) transcripts from the total nuclear RNA pool into a steady-state population of polyadenylated RNA, and that this accounts for the increased expression of the IVS(+) genes. It has been suggested that genes which do not require an intron for efficient expression may have strong signals for, and/or alternative methods of, 3' end processing.

How might the presence of an intron increase the efficiency of polyadenylation? It has been suggested that the intron sequesters snRNP particles (and/or other factors) which may be involved in splicing and 3' processing (Huang and Gorman, 1990). There is some evidence to support such a proposition. *In vitro* complementation studies of 3' end processing is inhibited by treatment of the extract with micrococcal nuclease and with an anti-Sm antibody (Hashimoto and Steitz, 1986; Gilmartin *et al.*, 1988). These observations suggest that snRNPs are involved in 3' end processing. In cross-linking studies the hnRNP C protein (thought to be involved in splicing) is found to be associated with the polyadenylation signal sequence (Moore *et al.*, 1988). Two other proteins, one of which possesses an Sm-epitope, are also found to be in direct contact with the polyadenylation site.

Pre-mRNA splicing appears to take place on the nuclear matrix (Ciejeck *et al.*, 1982), indeed pre-mRNA transcripts and the splicing intermediates are found associated with it (Zeitlin *et al.*, 1987). In a recent study the pre-mRNA of the acetylcholine receptor gene has been localised to specific regions of the nuclear membrane which is in contrast to the random distribution of the actin gene pre-mRNA (Berman *et al.*, 1990). It has been proposed that the nuclear membrane regulates the export of certain RNA molecules and it is conceivable that transcripts may be spliced at specific regions on the nuclear membrane as a means of controlling their expression.

Introns can contain elements involved in the control of transcription (Aronow *et al.*, 1989 and references therein). The mouse ribosomal protein gene *RPL32* (Chung and Perry, 1989) and the human adenosine deaminase gene (*ADA*) (Aronow *et al.*, 1989), both contain transcriptional activators in their first introns. Indeed it is interesting to note that increases in gene expression appear to be greater when an intron is placed at the 5' end of a transcript (Callis *et al.*, 1987). This effect need not necessarily be a result of intron sequences involved in transcriptional control. It could be related to the early sequestration of factors required for splicing and/or 3' processing (see above). Introns can produce negative effects on gene expression. The presence of intron-4 in the transcript of the human *PCNA* gene reduces the level of mRNA and indeed is required for the correct regulation of gene expression (Ottavio *et al.*, 1990). It is possible that the removal of intron-4 of *PCNA* is regulated, indeed the *RPL32* transcript of yeast can regulate the splicing of its own transcript (Dabeva *et al.*, 1986). Recently a promoter for antisense RNA was identified in the yeast actin gene intron, though the function of the antisense transcript is unknown (Thompson-Jager and Domdey, 1990).

## 1.12 INTRONS AND EVOLUTION

There are two main theories advocated to account for the presence of introns in genes. The first suggests that introns were present in the original nucleic acid molecules that constituted the so-called "RNA world" (Gilbert, 1986). This idea is an extension of the theory that introns pre-dated the divergence of prokaryotic and eukaryotic organisms (Doolittle, 1978; Darnell, 1978). In this scenario introns were subsequently lost from prokaryotic genes as a consequence of the high selective pressure to minimise the size of the genome to enable rapid DNA replication. In contrast, introns were maintained in eukaryotic genes because they facilitated diversity and rapid evolution and enabled the control of gene expression. The second theory suggests that introns were inserted into the genes of eukaryotes after the divergence from prokaryotes (Crick, 1978; Orgel and Crick, 1980).

There is now a considerable amount of evidence supporting the former theory. Introns

have been observed in prokaryotic organisms (Schmidt, 1985) and studies of chloroplast and nuclear genes has indicated that introns were present before the divergence of prokaryotes and eukaryotes (Shih *et al.*, 1988). If introns were present in ancestral genes then the variability in intron content exhibited by functionally similar genes in different species, should be explainable by the specific loss of introns, and indeed this phenomenon has been documented (Perler *et al.*, 1980). However there is also evidence to suggest that introns were inserted into eukaryotic genes after the divergence of eukaryotes from prokaryotes (Sogin, *et al.*, 1986; Dibb and Newman, 1989; Tani and Ohshima, 1989). Group I introns can insert into specific regions of DNA (reviewed by Dujon, 1989; Lambowitz, 1989) and indeed Group II introns (and some Group I introns) may insert into DNA via an RNA intermediate (Woodson and Cech, 1989; Morl and Schmelzer, 1990; Augustin *et al.*, 1990). The relationship of the Group I and II introns to mobile genetic elements is an interesting one (see Scazzocchio, 1989) and there are reports of pre-mRNA-type introns possessing sequence homology to transposons (Tournier-Lasserre, 1989). These observations suggest that present day genomes are in flux, subject to intron deletion and insertion.

It has been postulated that RNA was the first replicator molecule in pre-biotic evolution and in the absence of proteins was both an informational and functional molecule (Sharp, 1985; Gilbert, 1986; Lamond and Gibson, 1990). It was hypothesised that introns could recombine into other RNA molecules and in the process "carry" exon sequences with them. This recombination event would increase the diversity of molecules by generating novel combinations of exons. The discovery that Group I and II introns could self-splice and re-integrate into other RNA molecules lent great weight to this hypothesis (Cech and Bass, 1986; Woodson and Cech, 1989; Augustin *et al.*, 1990; Morl and Schmelzer, 1990). The mechanism by which Group I, II and pre-mRNA-type introns are removed is fundamentally similar (Sharp *et al.*, 1985; Cech and Bass, 1986) and thus it is thought that pre-mRNA type introns arose from the ancestral self-splicing introns. Based on this possible connection it is proposed that the catalytic activity involved in pre-mRNA splicing resides with the snRNA components of the spliceosome (Brow and Guthrie, 1989). The identification of other catalytic functions of RNA, e.g. from an RNA



polymerase to a sequence-specific RNA endonuclease, has supported the arguments for its role in the earliest stages of pre-biotic evolution (reviewed by Lamond and Gibson, 1990). Indeed Group I self-splicing introns can specifically bind L-arginine, suggesting their involvement in the creation of the translational apparatus (Hicke *et al.*, 1989).

What would be the selective pressure for the retention of introns? As previously indicated (see Section 1.11) introns can contain transcriptional regulatory elements and be involved in other methods of regulating gene expression e.g. alternative splicing. Indeed pre-mRNA splicing may be linked to other RNA processing events such as polyadenylation and transport. Introns themselves can encode proteins (Dujon, 1989). Another postulated pressure for the retention of introns is that of the process known as "exon-shuffling" (Gilbert, 1978). Non-homologous recombination within introns may serve to generate coding sequence diversity during evolution. In this scenario exon sequences correspond to functional domains of proteins (Blake, 1978), thus "exon-shuffling" leads to the reassortment of protein domains (Gilbert, 1985). Indeed genes made up of exons that encode specific protein domains found in other genes have been identified (Sudhof *et al.*, 1985a, 1985b). It has been proposed that "exon-shuffling" only became prevalent after the divergence of eukaryotes and prokaryotes (Patthy, 1987).

### 1.13 THIS THESIS

In this thesis I present the following;

- (i) the creation of a number of plasmid constructs which enabled the development of a novel screening procedure for conditional-lethal *prp* mutants in *Saccharomyces cerevisiae*,
- (ii) the development of a colony screening method which positively identifies *prp* mutants,
- (iii) the isolation and initial characterisation of temperature-sensitive *prp* mutants.

In addition I include an Appendix section which briefly describes a set of experiments designed to investigate the process of commitment of a pre-mRNA transcript to splicing.



## CHAPTER 2

### MATERIALS AND METHODS

#### 2.1 MATERIALS

##### 2.1.1 Suppliers of Laboratory Reagents

Restriction enzymes and other DNA modifying enzymes:

Amersham International, Bethesda Research Laboratories, Boehringer Mannheim, New England Biolabs and Pharmacia.

Radiochemicals:

Amersham International.

Deoxynucleotides, dideoxynucleotides and ribonucleotides:

Pharmacia.

Acrylamide and NN'-methylene bisacrylamide:

BDH Chemicals, "Electran" grade.

BSA and amino acids:

Sigma.

Ultra-pure agarose:

Bethesda Research Laboratories.

Media reagents:

Difco Labs.

Standard laboratory reagents (analytical grade or better):

BDH, Fisons, Sigma, Bio-Rad and Serva.

### 2.1.2 Growth Media

Unless otherwise stated all quantities given are for 1 litre volumes. These solutions were autoclaved (see Section 2.2.1.2) and stored at room temperature. Amino acids were filter sterilised (see Section 2.2.1.2) and along with antibiotics were stored as concentrated solutions and added after autoclaving the media. Vitamin-free Casamino acids (Difco) and adenine sulphate were added prior to sterilisation. When Casamino acids were used, tryptophan was added after autoclaving and cooling the media. Selective growth for *URA3*-containing plasmids still occurs in the presence of vitamin-free Casamino acids.

#### 2.1.2.1 Bacterial Media

Luria Broth (LB):	10g bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, pH to 7.2 with NaOH.
Luria agar (LB-agar):	LB plus 15g agar (Difco).
LB-ampicillin:	LB plus 0.1mg/ml ampicillin.
LB-amp agar:	LB-agar plus 0.1mg/ml ampicillin.
BBL agar:	10g trypticase (Baltimore Biological Laboratories), 5g NaCl, 10g agar (Difco).
BBL top agar:	as for BBL agar except 6.5g agar (Difco).
M9 salts x4:	28g Na <sub>2</sub> HPO <sub>4</sub> , 12g KH <sub>2</sub> PO <sub>4</sub> , 2g NaCl, 4g NH <sub>4</sub> Cl.
M9 medium (400ml):	100ml M9 x 4, 0.4ml 1M MgSO <sub>4</sub> , 0.8g glucose, amino acids and/or other requirements to 20µg/ml.
M9 agar (400ml):	M9 medium plus 300ml water agar (Difco).

#### 2.1.2.2 Yeast Media

YPDA:	10g bacto yeast extract (Difco), 20g bacto-peptone (Difco), 20g glucose, 20mg adenine sulphate.
YPDA agar:	YPDA plus 20g agar (Difco).
YMG/Cas:	6.7g yeast nitrogen base w/o amino acids (Difco), 20mg adenine

sulphate, 20g glucose, 10g vitamin-free Casamino acids (Difco) and 20mg tryptophan.

YMG/Cas agar: YMG/Cas plus 20g agar.

YMM: 6.7g yeast nitrogen base w/o amino acids (Difco), 20g glucose, 20mg relevant amino acid or base.

YMM agar: YMM plus 20g agar (Difco).

YMR/Cas: YMG/Cas except 20g raffinose instead of glucose.

YMR/Gal/Cas agar: MR/Cas plus 20g galactose and 20g agar (Difco).

YMR/G/Cas: YMR/Cas plus 500mg glucose.

Sporulation agar: 10g Potassium acetate, 1g bacto yeast extract (Difco), 0.5g glucose, 20g agar (Difco), 20mg each relevant amino acid and/or adenine sulphate.

#### 2.1.3 Bacterial Strains

All strains of *E.coli* used in this work are indicated in Table 2.1.

#### 2.1.4 Yeast Strains

All strains of *Saccharomyces cerevisiae* used in this work are indicated in Table 2.2.

TABLE 2.1 BACTERIAL STRAINS

strain	genotype	Reference
BMH71-18 <i>mutL</i>	$\Delta(lac, pro)$ , <i>supE</i> , <i>thi</i> <sup>-</sup> , <i>mutL</i> , <i>F'(proAB, lacI, lacZΔM15)</i> .	Kramer <i>et al.</i> , (1984)
HB101	<i>F</i> <sup>-</sup> , <i>hsdS20</i> ( <i>r</i> <sup>-</sup> <i>B</i> , <i>m</i> <sup>-</sup> <i>B</i> ), <i>recA13</i> , <i>ara-14</i> , <i>proA2</i> , <i>lacY1</i> , <i>galK2</i> , <i>rpsL20</i> ( <i>Sm</i> <sup>r</sup> ), <i>xyl-5</i> , <i>mtl-1</i> , <i>supE44</i> .	Boyer and Roulland- Dussoix. (1969)
JM101	$\Delta(lac, pro)$ , <i>codA</i> , <i>codB</i> , <i>trmB</i> , <i>argF</i> , <i>F'(traD36, proAB,</i> <i>lacI<sup>q</sup>, lacZΔM15)</i> .	Messing <i>et al.</i> , (1983)
NM522	<i>hsdΔ5</i> , ( <i>lac, pro</i> ), <i>F'(lacI<sup>q</sup>, lacZΔM15, pro+)</i> .	N. Murray. (this dept.)

TABLE 2.2 *SACCHAROMYCES CEREVISIAE* STRAINS

strain	genotype	source
KY117	<i>MAT<math>\alpha</math></i> , <i>his3-<math>\Delta</math>200</i> , <i>lys2-801</i> , <i>ade2-101</i> , <i>trp1-<math>\Delta</math>1</i> , <i>ura3-52</i> .	K. Struhl. (Harvard)
KY118	<i>MAT<math>\alpha</math></i> , as for KY117.	"
SPJ0.9	<i>MAT<math>\alpha</math></i> , <i>lys2</i> , <i>leu2</i> , <i>his3-<math>\Delta</math>1</i> , <i>ura3-52</i> .	S. Jackson. (this lab.)
DBY746	<i>MAT<math>\alpha</math></i> , <i>leu2-3 and 112</i> , <i>his3-<math>\Delta</math>1</i> , <i>trp1-289</i> , <i>ura3-52</i> .	D. Botstein. (M.I.T.)
S150-2B 2 $\mu$	<i>MAT<math>\alpha</math></i> , <i>leu2-3 and 112</i> , <i>his3-<math>\Delta</math>1</i> , <i>trp1-289</i> , <i>ura3-52</i> .	Wellcome.
DBY745	<i>MAT<math>\alpha</math></i> <i>leu2</i> , <i>ade1-100</i> , <i>ura3-52</i> .	Legrain and Rosbash (1989).
DJY72	<i>MAT<math>\alpha</math></i> , <i>prp1-1</i> , <i>leu2-3 and 112</i> , <i>his3-<math>\Delta</math>1</i> , <i>trp1-289</i> , <i>ade</i> , <i>ura3-52</i> .	D. Jamieson. (this lab.)
DJY73	as for DJY72.	"

DJY74	<i>MAT<math>\alpha</math>, prp1-1, leu2-3 and 112, his3-<math>\Delta</math>1, trp1-289, ura3-52, ura1.</i>	"
DJY36	<i>MATa, prp2-1, ade, ura3-52.</i>	"
SPJ2.53	<i>MAT<math>\alpha</math>, prp2-1, leu2, lys2, his, ade, tyr1, ura3-52.</i>	S. Jackson. (this lab.)
SPJ2.54	as for SPJ2.53.	"
SPJ3.33	<i>MATa, prp3-1, leu2, lys2, his, ura3-52.</i>	S. Jackson. (this lab.)
SPJ3.38	<i>MAT<math>\alpha</math>, as for SPJ3.33.</i>	"
SPJ4.41	<i>MATa, prp4-1, leu2, ura3-52.</i>	"
SPJ4.4	<i>MAT<math>\alpha</math>, prp4-1, leu2, lys2, his7, ade, ura3-52.</i>	"
SPJ5.41	<i>MAT<math>\alpha</math>, prp5-1, leu2, his, ura3-52.</i>	"
J89	<i>MAT<math>\alpha</math>, prp5-1, ade, trp, arg.</i>	J. Warner. (New York)
J10	<i>MATa, prp5-1, lys2, his7, ade1, ade2, tyr1, gall, ura1.</i>	"

SPJ6.66	<i>MATa, prp6-1, lys2, his, ade,</i> <i>ura3-52.</i>	S. Jackson. (this lab.)
SPJ6.68	<i>MAT<math>\alpha</math>, prp6-1, lys2, his, ura3-52.</i>	"
SPJ7.6	<i>MAT<math>\alpha</math>, prp7-1, leu2, lys2, his.</i>	"
SB	<i>MATa, prp7-1, lys2, his7, ade1,</i> <i>ade2, tyr1, ura1.</i>	A. Lustig. (New York)
AH	<i>MAT<math>\alpha</math>, prp7, leu, ade1+/-2, ura.</i>	"
86	<i>MATa, prp7-1, lys2, his7, ade1,</i> <i>ade2, tyr1, ura1, mal, gal.</i>	L. Hartwell. (1970)
SPJ8.31	<i>MATa, prp8-1, leu2, his3, ura3-52.</i>	S. Jackson. (this lab.)
SPJ8.3	<i>MAT<math>\alpha</math>, prp8-1, leu2, ade, ura3-52.</i>	"
SPJ8.36	<i>MAT<math>\alpha</math>, prp8-1, leu2, his3 and 7,</i> <i>ade, ura3-52.</i>	S. Jackson. (this lab.)
J17	<i>MATa, prp9-1, lys2, his7, ade1,</i> <i>ade2, tyr1, ura1, gal1.</i>	J. Warner. (New York)

J93	<i>MAT<math>\alpha</math>, prp9-1, leu, ade, trp,</i> <i>arg, ura.</i>	"
A635	<i>MATa, prp11-1, leu2, his4-512,</i> <i>tyr1, ura3-52.</i>	J. Haber. (Brandeis)
SPJ11.6	<i>MAT<math>\alpha</math>, prp11-1, leu2, his4-512,</i> <i>ade, tyr1, ura3-52.</i>	S. Jackson. (this lab.)
60-2a	<i>MATa, prp17, lys2-801, his3-<math>\Delta</math>200,</i> <i>ade2-101, ura3-52.</i>	J. Abelson. (Cal. Tec.)
20-1b	<i>MAT<math>\alpha</math>, as for 60-2a.</i>	"
27-5c	<i>MATa, prp18, his3-<math>\Delta</math>200, ade2-101,</i> <i>tyr1, ura3-52.</i>	"
27-6c	<i>MAT<math>\alpha</math>, prp18, lys2-801, his3-<math>\Delta</math>200,</i> <i>ade2-101, ura3-52.</i>	"
10A-6B	<i>MATa, prp19, leu2, lys2-801,</i> <i>his3-<math>\Delta</math>200, ade2-101, ura3-52.</i>	"
10A-3D	<i>MAT<math>\alpha</math>, as for 10A-6B.</i>	"
20-6D	<i>MATa, prp20, his3-<math>\Delta</math>200, ade2-101,</i> <i>tyr1, ura3-52.</i>	"



20-6A	<i>MAT<math>\alpha</math></i> , as for 20-6D.	"
4C	<i>MATa</i> , <i>prp21</i> , as for 20-6D.	"
2A	<i>MAT<math>\alpha</math></i> , <i>prp21</i> , as for 60-2a.	"
9-6C	<i>MATa</i> , <i>prp22</i> , as for 20-6D.	"
9-4A	<i>MAT<math>\alpha</math></i> , <i>prp22</i> , as for 60-2a.	"
29-5B	<i>MATa</i> , <i>prp23</i> , as for 27-5C.	"
29-6A	<i>MAT<math>\alpha</math></i> , <i>prp23</i> , <i>lys2-801</i> , <i>his3-<math>\Delta</math>200</i> , <i>ade2-101</i> , <i>tyr1</i> , <i>ura3-52</i> .	"
40-5C	<i>MATa</i> , <i>prp24</i> , as for 60-2a.	"
35-3a	<i>MAT<math>\alpha</math></i> , <i>prp24</i> , as for 60-2a, ( <i>ura3?</i> ).	"
23-2C-1C	<i>MATa</i> , <i>prp25</i> , as for 27-5C.	"
23-2C-1A	<i>MAT<math>\alpha</math></i> , <i>prp25</i> , as for 60-2a.	"
D-5A-3A	<i>MATa</i> , <i>prp26</i> , as for 27-5C.	"
40-1a	<i>MATa</i> , <i>prp27</i> , as for 60-2a.	"

35-3b

*MAT $\alpha$* , *prp27*, as for 60-2a.

"

DJY63

*MAT $\alpha$* , *SPP81-3*, *leu2*, *his3*,  
*ura3-52*.

D. Jamieson.

(this lab.)

DJY65

*MAT $\alpha$* , *prp8-1*, *SPP81-3*, *leu2*,  
*his3*, *ura3-52*.

"



### 2.1.5 General Plasmids and Bacteriophage Vectors

Plasmids used for the isolation of gene-specific DNA fragments (utilised as radioactively labelled probes in Northern blot analysis) and as general vectors in cloning procedures are described in Table 2.3 and in Fig. 2.1. All DNA fragments used as probes were gel purified (see Section 2.2.5.7) and radiolabelled as described in Section 2.2.5.10.

### 2.1.6 Synthetic Oligodeoxynucleotides

Table 2.4 provides the sequences of the oligodeoxynucleotides used in primer extension studies (see Section 2.2.8.5 and Chapter 4.1.1). All oligodeoxynucleotides were synthesised and HPLC purified by the Oswel DNA Service based in the department of Chemistry at the University of Edinburgh (including those oligodeoxynucleotides described in Chapter 3). All  $T_m$  values were calculated by the "Wallace rule" i.e.  $T_m = 4(G+C \text{ content}) + 2(A+T \text{ content})$ ; (Suggs *et al.*, 1981). The concentration of the oligodeoxynucleotide-containing solution was determined using the following formula: 0.2  $\mu$ moles of a 20mer produces an O.D. 260<sub>nm</sub> of 20 units and is equivalent to 1mg of DNA (Oswel DNA Service). Oligodeoxynucleotide integrity was checked by phosphorylating the 5' end with [ $\gamma$ -32P]rATP (employing T4 polynucleotide kinase; see Section 2.2.5.12), and analysed by polyacrylamide gel electrophoresis (see Section 2.2.5.12). By comparing the sizes and patterns of radioactive signal with the M13 universal sequencing primer (17mer), those oligonucleotides that were of the wrong size and/or weakly labelled were quickly detected and discarded (results not presented).

TABLE 2.3 GENERAL PLASMIDS AND BACTERIOPHAGE VECTORS

Vector	Description	Reference
M13mp9	<i>E.coli</i> M13-bacteriophage based	Messing.J.
M13mp18	vectors used in SDM and DNA sequencing.	(1983)
pSPT19	<i>E.coli</i> plasmids containing a multiple cloning site, used as a host vector for the lacZ gene.	Pharmacia (U.K.)
pYA301	A 1038bp <i>Bam</i> HI/ <i>Bgl</i> III fragment from this plasmid, containing E1, IVS-1 and part of E2 of the actin gene, was used as a probe in Northern blots.	Gallwitz and Sures (1980)
pBRCYH-3	A 2kb <i>Eco</i> RI fragment of YEpCYH-1 containing 5' flanking sequences, E1, IVS-1 and part of E2 of the CYH2 gene was cloned into pBR322. This fragment is present in 3 copies (hence 3) and was used as a probe in Northern blots.	Made by G. Anderson. (this lab.)
pUC-GPM	pUC19 based vector containing the yeast phosphoglycerate mutase gene (GPM), which was removed as a 1.3 kb <i>Hind</i> III/ <i>Sal</i> I fragment and used as a probe in Northern blots.	M. White. (L. Gilmore's lab. Dept. of Biochem. Edinburgh.

## FIGURE 2.1: GENERAL PLASMID AND PHAGE VECTORS

The plasmids and phage used in general cloning experiments, and those used for the isolation of specific gene fragments, are indicated along with their sizes and relevant restriction sites. The plasmids are not drawn to scale.

*Bam*HI = B

*Eco*RI = E

*Ssp*I = Ssp

*Bgl*II = Bg

*Hind*III = H

*Sal*I = S

Figure 2.1

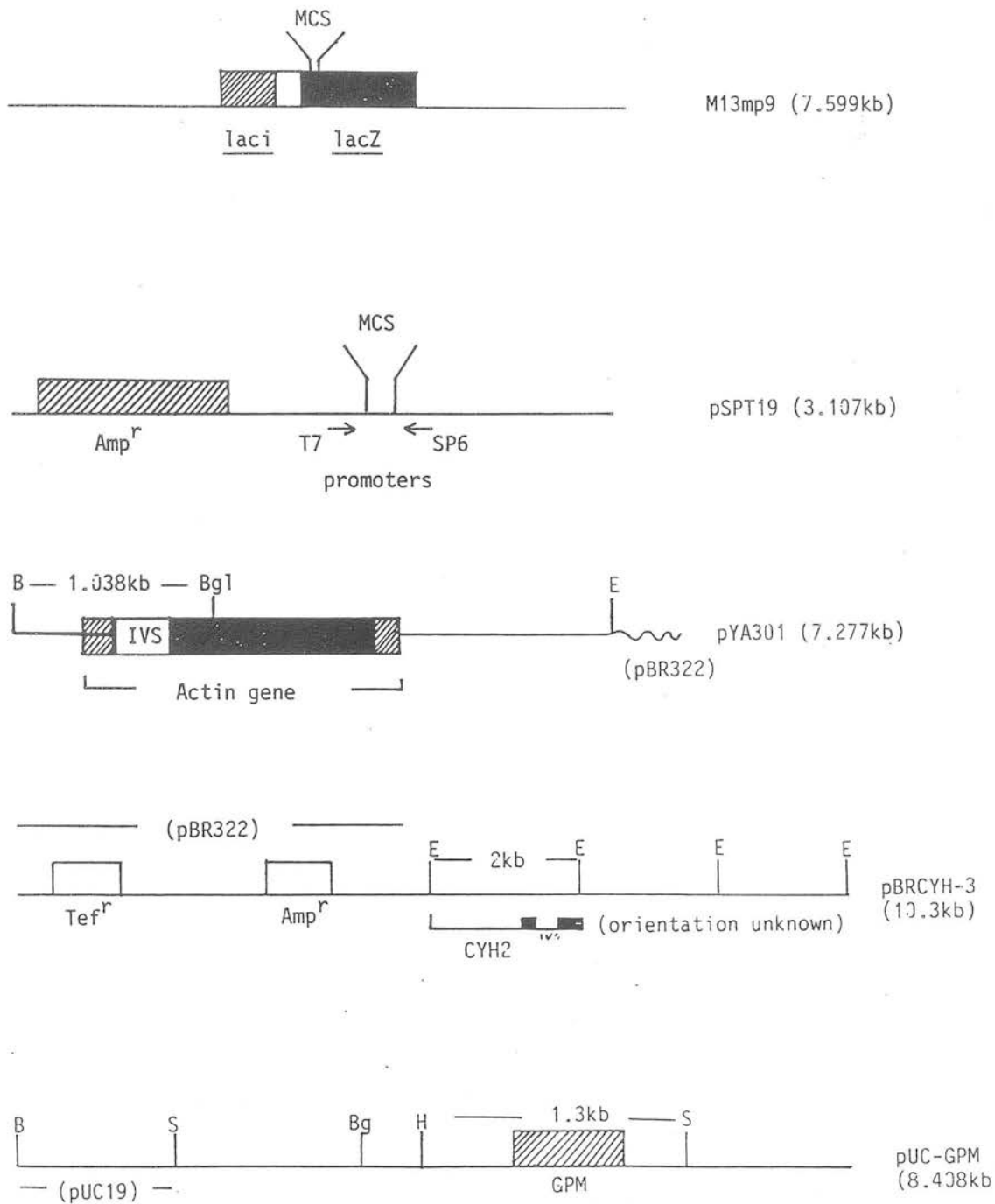


TABLE 2.4 SYNTHETIC OLIGODEOXYNUCLEOTIDES

Name	Oligodeoxynucleotide Sequence and Description
------	---

083C	5'-CTGCAAGGCGATTAAGTTGG-3' 20mer, T <sub>m</sub> = 60°C
------	---

This oligodeoxynucleotide is complementary to the *lacZ* transcript, 47 bases downstream of the *Bam*HI site present at the 5' end of the coding sequence (see Chapter 3.3.1). Thus in all *MATa1* fusion constructs described in Chapter 3 this oligodeoxynucleotide was able to differentiate between the pre-mRNA and mRNA generated from each plasmid (it should be noted that there is no endogenous *lacZ* gene in any of the yeast strains tested). Using this oligodeoxynucleotide the sizes of primer extension products generated from the pre- and mRNA encoded by pJBM-4 are 381nts and 322nts respectively. The IVS-E2 lariat intermediate gives rise to a primer extension product of 155nts.

JMO-PGK	5'-GGGTGGTGTTCCTCAAACG-3' 18mer, T <sub>m</sub> = 56°C
---------	--

(050B)

This oligodeoxynucleotide was complementary to the mRNA sequence located 200 nucleotides downstream of the transcription start site of the intronless yeast gene phosphoglycerate kinase (PGK), (Hitzeman, 1982). It was used as a control for changes in RNA levels such that any detectable increases in pre-mRNA could be distinguished from changes in RNA not subject to splicing.

JMACT1	5'-TCAATAACCAAAGCAGCAAC-3' 20mer, T <sub>m</sub> = 56°C
--------	---

(203D)

This oligodeoxynucleotide is complementary to the E2 region of the actin mRNA that is 2bp downstream of the 3' splice site (Ng and Abelson, 1980). Thus it will detect pre- and mRNA and the IVS-E2 lariat intermediate. The sizes of the DNA species generated by primer extension using this oligodeoxynucleotide are 462nts, 153nts and 65nts respectively.

JMACT2      5'-GAAATCTCTCGAGCAATTGG-3' 20mer, T<sub>m</sub> = 58°C

(204D)

This is complementary to the actin transcript 54bp downstream of the 5' splice site and 186bp upstream of the TACTAAC box. It therefore detects the IVS-E2 lariat intermediate and/or the IVS lariat product (it cannot distinguish between them as both produce extension products of 74nts), and pre-mRNA (205nts extension product).

JMCYH2A      5'-TTACCGATACGACCTTTACC-3' 20mer, T<sub>m</sub> = 58°C

(205A)

This is complementary to the E2 region of the ribosomal protein gene *CYH2* transcript which is 2bp downstream of the 3' splice site and thus binds to pre- and mRNA, and IVS-E2 lariat intermediate (Kaufer *et al.*, 1983). This gene has at least 3 transcription start sites and this is reflected in the sizes of the primer extension products: pre-mRNA, 616, 612 and 606nts; mRNA, 106, 102 and 96nts; IVS-E2 lariat, 66nts.

JMCYH2B      5'-ATATTGGTTGCACAACACTA-3' 20mer, T<sub>m</sub> = 54°C

(206D)

This is complementary to the region of the *CYH2* intron that is 75nts downstream of the 5' splice site and 385nts upstream of the TACTAAC box. It has the same recognition pattern as JMACT2 and produces primer extension products of lengths 160, 156 and 150nts for pre-mRNA and 75nts for IVS-E2 lariat and/or IVS lariat product.



## 2.2 METHODS

### 2.2.1 General Guidelines

#### 2.2.1.1 Centrifugation

Unless indicated all procedures were performed at room temperature. Centrifugation of volumes greater than 50ml was achieved using a Sorvall GS-3 rotor at 4,100g (5,000rpm); volumes between 50 and 1.5ml were centrifuged in either a Sorvall SS34 or HB4 rotor at 16,000g (11,500 and 10,000rpm respectively), or in a IEC CENTRA-4X bench-top centrifuge (with swing out rotor) at 5000rpm; and volumes less than or equal to 1.5ml in a microcentrifuge at 17,500g (12,000rpm).

#### 2.2.1.2 Sterilization of Solutions

When necessary solutions were sterilized prior to use by placing them in an autoclave (120°C, 15psi), for 20 minutes. Constituents of solutions which could not be treated in this manner e.g. SDS, tryptophan, DTT,  $\beta$ -mercaptoethanol and antibiotics etc., were added after this process. Such solutions were sterilized by passage through a disposable acrodisc filter-sterilizing assembly (0.45 $\mu$ m pore-size; Gelman Sciences).

#### 2.2.1.3 Deionisation of Solutions

Solutions such as formamide, acrylamide and glyoxal were deionised by mixing with 0.1-0.2 volumes of analytical grade mixed-bed resin (20-50 mesh, Bio-Rad Laboratories), for 30 minutes. The resin was then removed by filtration through Whatman No.1 filter paper. Glyoxal was stored in tightly sealed eppendorf tubes, in 100 $\mu$ l aliquots, at -70°C. Acrylamide and formamide were stored at 4°C

#### 2.2.1.4 Autoradiography

[<sup>32</sup>P]-labelled nucleic acids were detected by exposure to preflashed X-ray film (Cronex) at -70°C in the presence of a calcium-tungstate intensifying screen (Amersham Review 23, Laskey, R. A.).

[<sup>35</sup>S]-labelled DNA fragments were detected by exposure to unflashed X-ray film (Cronex) at room temperature without an intensifying screen.

#### 2.2.1.5 Photography

Photographs of agarose gels containing nucleic acids stained with 1µg/ml ethidium bromide were obtained by visualising the nucleic acid with a short-wave length UV (254<sub>nm</sub>) transilluminator, and then exposing the gel to Polaroid 667 professional film.

### 2.2.2 Bacteria and Phage: General Microbiological Methods

#### 2.2.2.1 Propagation and Storage

*E.coli* cells were grown at 37°C. Liquid cultures were shaken continuously, and growth was monitored by determination of the cell density, measured at O.D. 550<sub>nm</sub>. *E.coli* strains HB101 and DH5 were propagated on LB-agar whereas JM101, NM522 and BMH71-18 *mutL* were maintained on M9-agar to select for the F' episome. Transformed *E.coli* strains were grown on LB-amp to select for the plasmid of interest. Permanent *E.coli* stocks (85% (v/v) stationary culture grown in the relevant medium) were stored at -70°C in 15 % (v/v) glycerol.

#### 2.2.2.2 Preparation of Competent Cells

A 1% (v/v) inoculum of a stationary culture in 100ml LB, was shaken vigorously until

it reached an O.D. 550<sub>nm</sub> of roughly 0.5. After chilling on ice for 10 minutes the cells were sedimented (4,100g, 5 minutes at 4°C), resuspended in 50% (original volume) cold 75mM CaCl<sub>2</sub> and incubated on ice for 20 minutes. The cells were re-sedimented (4,100g, 5 minutes at 4°C, resuspended in 4% (original volume) cold TFBII (10mM MOPS pH6.8, 10mM RbCl, 75mM CaCl<sub>2</sub> 15% (v/v) glycerol) and incubated on ice for 20 minutes. They were then aliquoted (600μl) into eppendorf tubes and either used immediately or snap frozen in liquid nitrogen and stored at -70°C.

### 2.2.2.3 Transformation

To a glass test-tube containing 200μl of competent cells, roughly 50ng of plasmid DNA (in a volume not exceeding 100μl) was added and thoroughly mixed. After leaving on ice for 20 minutes the mixture was incubated at 42°C for 2.5 minutes (with occasional mixing), and then replaced on ice for a further 5 minutes. It was then transferred to an eppendorf tube containing 1ml of LB and incubated on a rotating wheel at 37°C for 40-60 minutes. Aliquots of cells were then spread onto L-amp plates to select for cells containing the plasmid.

### 2.2.2.4 Isolation of M13 Plaques

Single-stranded or RF M13 DNA was introduced into *E.coli* strains as described in Section 2.2.2.3. During the transformation procedure BBL top agar was melted, aliquoted (3.5ml) into small 5ml glass test-tubes and incubated at 42°C for 5 minutes followed by the addition of 30μl X-gal (20mg/ml) and 20μl IPTG (24mg/ml). Aliquots of the transformed cells (2-100μl) were then mixed with 200-300μl plating cells (JM101 or NM522), which could be stationary or freshly grown cultures (OD 550nm of 0.2-0.5). After this, these cells were added to the test-tubes containing the BBL top agar, X-gal and IPTG, mixed and poured over the surface of a BBL plate. The top agar was allowed to set and then the plates were incubated overnight at 37°C.

## 2.2.3 Yeast: General Microbiological and Genetic Methods

### 2.2.3.1 Propagation and Storage

*prp* strains were grown at 23°C while wild-type strains were propagated at 30°C. Liquid cultures were continuously shaken and cell density was determined by measuring the O.D.<sub>600nm</sub>. Untransformed strains were grown on YPDA, whereas selective growth for strains harbouring plasmids containing the *URA3* gene was achieved by growth on YMG/Cas. Strain genotypes were determined by screening for growth on YMM medium supplemented with the relevant amino acids and bases. All plates were stored at 4°C. Permanent yeast stocks were maintained at -70°C; 75% (v/v) yeast stationary culture grown in the relevant medium added to 25% (v/v) glycerol. In order to recover a higher percentage of viable yeast cells from the frozen stocks, 4ml of stationary culture was sedimented and resuspended in 0.75ml of YPDA or YMG/Cas and this was added to the glycerol.

### 2.2.3.2 Generation of Diploid Strains

Cells of opposite mating type were mixed in 30µl of dH<sub>2</sub>O on the relevant medium and incubated at 23°C or 30°C for 2-3 days (during which time mating took place). Cells from the mixture were then spread onto YMM plates supplemented with those amino acids and bases which allowed growth of the diploid strain only. If required, single colonies were isolated by streaking the cells.

### 2.2.3.3 Sporulation and Spore Recovery

Diploid cells were mixed in 30µl of dH<sub>2</sub>O on sporulation plates and incubated at 23°C or 30°C for 3-7 days. They were then suspended in 90µl of Spore buffer (1.2M sorbitol, 10mM EDTA, 100mM citric acid) until the solution was slightly turbid, and 10µl of β-Glucuronidase

(Boehringer Mannheim) was added to digest the asci walls. The suspension was gently mixed and left at room temperature for 30-45 minutes. An aliquot of this mixture was spread onto YPDA plates and the Singer micro-manipulator was used to isolate single asci and to dissect them into individual spores. These plates were incubated at 23°C or 30°C until individual spores had germinated and grown.

#### 2.2.3.4 Mating-Type Analysis

The mating-type of relevant strains was determined by mixing them on YPDA with those of known mating-type. These were left to incubate at the relevant temperature for 2-3 days, during which time mating did or did not occur. The cells were then replica-plated onto YMM supplemented with those amino acids and bases which allowed growth of the diploid cells only, and incubated at the relevant temperature for 3-7 days.

#### 2.2.3.5 Complementation Analysis

To determine whether different  $t^s$  mutants harboured  $t^s$  mutations in the same complementation group,  $t^s$  strains of opposite mating type were suspended in dH<sub>2</sub>O and mixed on YPDA plates. They were left to incubate at 23°C for 2-3 days, during which time mating took place. These were then replica-plated and incubated at 23°C and 36°C for 3-10 days. If the crosses generated diploids which failed to grow at 36°C, this indicated that the parent strains contained conditional lethal mutations in the same complementation group. The plates were incubated for at least 10 days to detect any possible growth at 36°C.

#### 2.2.3.6 Transformation

The lithium acetate method of transformation was used to introduce plasmid DNA into the appropriate yeast strains (Ito et al., 1983). A 1% (v/v) inoculum of a stationary culture

"TE" buffer	:	10mM Tris.HCl pH8 1mM EDTA
"TE" <sub>0.1</sub> buffer	:	10mM Tris.HCl pH8 0.1mM EDTA
5 X "P" buffer	:	1% (w/v) BSA 1% (w/v) Polyvinylpyrrolidone (PVP) 1% (w/v) Ficoll 250mM Tris.HCl pH7.5 0.5% sodium pyrophosphate 5% SDS
5 X Denhardt's solution	:	0.1% (w/v) BSA 0.1% (w/v) PVP 0.1% (w/v) Ficoll

---

was added to 50ml YPDA and grown until an O.D.  $600_{nm}$  of roughly 0.4 had been attained. The cells were harvested in 50ml falcon tubes (5 minutes, IEC CENTRA-4X bench-top centrifuge, swing-out rotor at 4,500rpm) and resuspended in 10ml sterile dH<sub>2</sub>O or TE. The cells were re-sedimented and resuspended in 5ml sterile dH<sub>2</sub>O or TE. After this 208 $\mu$ l of 2.5M lithium acetate (final concentration roughly 0.1M) was added and the tubes were incubated at the relevant temperature with gentle mixing for 1 hour. Aliquots of these cells (200 $\mu$ l) were transferred to eppendorf tubes and incubated with 1-10 $\mu$ g of DNA (in a volume of 100 $\mu$ l) at the appropriate temperature for 30 minutes (no shaking required). Following this incubation, 0.7ml of 50% PEG-4,000 (final concentration 35%) was added, mixed and the suspension left to incubate for a further hour. After heat-shocking for 5 minutes at 42°C the cells were harvested (17,500g, 5 minutes), washed in dH<sub>2</sub>O, re-sedimented (17,500g, 5 minutes) and finally resuspended in 100 $\mu$ l dH<sub>2</sub>O. The cells were then spread onto medium selecting for cells that had retained the introduced DNA.

#### 2.2.3.7 UV Mutagenesis and Isolation of *r<sup>s</sup>* Colonies

DBY745 and DBY746 transformed with pJBM-4 were grown to stationary phase in 20ml YMG/Cas. The cells were harvested (4,500rpm, IEC CENTRA-4X bench-top centrifuge, swing out rotor, 7 minutes), washed in 20ml sterile dH<sub>2</sub>O, sedimented (as before), resuspended in 20ml sterile dH<sub>2</sub>O and aliquoted into two sterile glass petri-dishes. All manipulations from this point were carried out either in the dark or with the cells shielded from light. The cell suspension was exposed to UV light (lids off; 10 ergs/mm<sup>2</sup>/sec. from a height of 32cm) for the required time to induce 80%-90% killing (see Chapter 5.1), then dispensed into universal bottles wrapped in tin-foil. The cells were sedimented by centrifugation (4,500, IEC Centra-4X bench-top centrifuge, swing-out rotor for 5 minutes), resuspended in 20ml YMG/Cas and aliquoted (200 $\mu$ l) into eppendorf tubes. The eppendorf tubes were placed into polystyrene trays which could be covered and taped onto a rotor. In this way the small aliquot volume used could be kept in the dark and still allow constant mixing to produce better aeration. After 3-5 days

incubation at 23°C each aliquot was diluted (usually 10<sup>-4</sup>), spread onto YMG/Cas agar and incubated at 23°C for 3-5 days. The colonies which grew were replica-plated onto YMG/Cas agar (incubated at 23°C and YMG/Cas agar containing 0.001% Phloxine B (incubated at 36°C). Phloxine B is a non-specific red-coloured dye which is thought to bind to intra-cellular factors. The rationale for its use is that in *t<sup>s</sup>* mutants the cell wall usually becomes "leaky" relatively quickly (because the cells are dying); thus Phloxine B penetrates these cells and as a consequence *t<sup>s</sup>* colonies exhibit a deep red colour. In contrast, wild-type colonies display a pink colour and against this background *t<sup>s</sup>* colonies are easier to detect. Protracted storage of cells on media containing this dye is not advisable as it results in cell death. This dye is used extensively with *Schizosaccharomyces pombe* (Fantes, P., pers. comm.). In *Saccharomyces cerevisiae* the amount of dye required for efficient detection of *t<sup>s</sup>* mutants may be lower than indicated as even wild-type colonies appeared pink; however 0.001% Phloxine B produced an acceptable comparison. Those colonies which appeared *t<sup>s</sup>* in this assay were isolated and re-screened for the *t<sup>s</sup>* phenotype on YMG/Cas agar. Colonies which still exhibited the *t<sup>s</sup>* phenotype were classed as *t<sup>s</sup>* mutants and assayed for potential *prp* defects as described in Chapter 4.

#### 2.2.3.8 UV Mutagenesis: Survival Curve

Aliquots of cells were removed at 30sec intervals during exposure to UV light (see Section 2.2.3.7), diluted (if necessary), spread on the relevant medium and incubated in the dark for 3-5 days. The number of colonies present at each time-point was counted and compared to a sample (from the same culture) which had not been subjected to UV light. In this way the percentage of cells which survived a specific exposure time was determined.



## 2.2.4 General Nucleic Acid Methods

### 2.2.4.1 Storage

Double-stranded DNA was stored at 4°C in dH<sub>2</sub>O or TE, single-stranded M13 DNA was stored at -20°C in TE0.1 and RNA was stored in RNase-free dH<sub>2</sub>O at -70°C.

### 2.2.4.2 Quantitation of Nucleic Acid

The concentration of a nucleic acid solution was determined by obtaining the O.D. 260<sub>nm</sub> reading (using a quartz cuvette in a LKB Biochrom Ultraspec II spectrophotometer), and estimating the concentration based on the assumption that a 50µg/ml solution of double-stranded nucleic acid and 40µg/ml solution of single-stranded nucleic acid has an O.D. 260<sub>nm</sub> of 1. Alternatively the concentration of double-stranded DNA was estimated by viewing the DNA in an agarose gel stained with ethidium bromide and comparing it with a known standard.

### 2.2.4.3 Precipitation With Ethanol

Nucleic acids were precipitated from aqueous solutions by the addition of 0.1 volume 3M sodium acetate pH5.5 and 2 or 2.5 volumes ethanol (previously incubated at -20°C for DNA and RNA samples respectively. The mixed samples were then placed on ice for roughly 15 minutes and sedimented by centrifugation in a microfuge or Sorvall SS34 or HB4 rotor (see Section 2.2.1.1). The samples were then washed by resuspension in an equal volume of 66% (v/v) ethanol (if RNA, DEP-treated dH<sub>2</sub>O was added to the ethanol; see Section 2.2.8.1), re-sedimented, dried in a vacuum desiccator and finally dissolved in the appropriate solvent.

### 2.2.4.4 Extraction With Phenol/Chloroform

"TE" buffer	:	10mM Tris.HCl pH8 1mM EDTA
"TE" <sub>0.1</sub> buffer	:	10mM Tris.HCl pH8 0.1mM EDTA
5 X "P" buffer	:	1% (w/v) BSA 1% (w/v) Polyvinylpyrrolidone (PVP) 1% (w/v) Ficoll 250mM Tris.HCl pH7.5 0.5% sodium pyrophosphate 5% SDS
5 X Denhardt's solution	:	0.1% (w/v) BSA 0.1% (w/v) PVP 0.1% (w/v) Ficoll

---

Nucleic acid solutions were deproteinised by vortexing vigorously and sequentially in an equal volume of phenol, phenol/chloroform and chloroform. The aqueous phase was separated and recovered each time by centrifugation in a microfuge (12,000g, 5 minutes) or a Sorvall SS34 rotor (4,100g, 7 minutes), usually at 4°C. After the final step the nucleic acid was precipitated as indicated in Section 2.2.4.3, which essentially removed any residual phenol. The phenol used in these procedures was equilibrated with either TE (double-stranded DNA), TE0.1 (single-stranded M13 DNA) or DEP treated ddH<sub>2</sub>O (RNA). The chloroform used was 24 parts chloroform to 1 part iso-amylalcohol.

## 2.2.5 DNA Methods

### 2.2.5.1 Plasmid DNA Isolation: Small Scale Preparations

The procedure used was based on the method of Birnboim and Doly (1979). LB-amp (5mls) was inoculated with a single transformed bacterial colony and incubated (with vigorous shaking) overnight at 37°C. An aliquot of the culture (1.5mls) was harvested (17,500g, 5 minutes) and the rest stored at 4°C. The cell pellet was suspended in 100µl of ice-cold solution I (50mM glucose, 10mM EDTA, 25mM Tris.HCl pH8; containing freshly added lysozyme to a final concentration of 4mg/ml) and incubated on ice for 30 minutes. After this 200µl of freshly prepared solution II (0.2N NaOH, 1% SDS) was added and the contents of the tubes gently mixed by inversion. Following a 5 minute incubation on ice, 150µl of ice-cold solution III (potassium acetate pH4.8; see below) was added, mixed and re-incubated on ice for 60 minutes. The cell debris was sedimented by centrifugation (17,500g, 5 minutes, 4°C), the supernate recovered, phenol/chloroform extracted (see Section 2.2.4.4) and ethanol precipitated (see Section 2.2.4.3). The precipitated nucleic acid was resuspended in 25µl of the desired buffer containing 20µg/ml RNase and 3-5µl aliquots were used in restriction enzyme analysis.

Solution III was made up as follows: 11.5ml glacial acetic acid and 28.5ml dH<sub>2</sub>O were

"TE" buffer	:	10mM Tris.HCl pH8 1mM EDTA
"TE" <sub>0.1</sub> buffer	:	10mM Tris.HCl pH8 0.1mM EDTA
5 X "P" buffer	:	1% (w/v) BSA 1% (w/v) Polyvinylpyrrolidone (PVP) 1% (w/v) Ficoll 250mM Tris.HCl pH7.5 0.5% sodium pyrophosphate 5% SDS
5 X Denhardt's solution	:	0.1% (w/v) BSA 0.1% (w/v) PVP 0.1% (w/v) Ficoll

---

added to 60ml of 5M potassium acetate. The resulting solution was 5M with respect to acetate and 3M with respect to potassium.

#### 2.2.5.2 Plasmid DNA Isolation: Midi-Preparations

In essence this method was a scaled-up version of that outlined in Section 2.2.5.1. It should be noted that only the final ethanol precipitation step in this protocol involves a 66% (v/v) ethanol wash. A single transformed bacterial colony was inoculated into 40mls LB-amp and incubated (with vigorous shaking) overnight at 37°C. The culture was harvested (IEC CENTRA-4X bench-top centrifuge, 4,500rpm for 7 minutes), suspended in 2mls solution I (see Section 2.2.5.1) and placed on ice for 30 minutes. After briefly vortexing, 4mls of solution II (see Section 2.2.5.1) was added, followed by further vigorous vortexing and incubation on ice (5 minutes). After this, 3ml of 3M sodium acetate pH5.5 was added and the were samples mixed by inversion and then incubated on ice for 30-60 minutes. Cell debris was sedimented by centrifugation (16,000g, 10-20 minutes, 4°C) and 16mls of cold ethanol was added to the recovered supernate, followed by incubation on ice for 15 minutes. The nucleic acid was sedimented by centrifugation (16,000g, 20 minutes, 4°C), dried (vacuum desiccator), suspended in 2mls of buffer B (0.1M sodium acetate pH5.5, 1mM EDTA, 0.1% SDS, 40mM Tris.HCl pH8), phenol/chloroform extracted and ethanol precipitated. The nucleic acid was once again sedimented by centrifugation (16,000g, 15 minutes) and dried and re-dissolved in 400µl TE to which 20µl 4M NaCl and 1ml ethanol was added. After ethanol precipitation, centrifugation (17,500g, 20 minutes, 4°C) and drying, the pellet was re-dissolved in 200ul TE (containing 0.1µg/µl RNase) and incubated at 37°C for 1 hour. Digestion of the RNA was followed by addition of 20µl 4M NaCl, phenol/chloroform extraction and ethanol precipitation. After sedimenting the nucleic acid (17,500g, 15 minutes), the pellet was finally washed with 66% (v/v) ethanol, dried (vacuum desiccator) and dissolved in 200µl TE. This procedure normally yielded 20-40µg of plasmid DNA.

#### 2.2.5.3 Plasmid DNA Isolation: Large Scale Preparation

Cells from a 500ml culture (1% (v/v) inoculum) incubated overnight, were harvested by centrifugation (4,100g, 15 minutes) and suspended in 6ml of sucrose mix (25% sucrose, 50mM Tris.HCl pH8, 40mM EDTA) containing 0.66mg/ml RNase. Lysozyme (10mgs; dissolved in 1ml sucrose mix) and 1ml 0.5M EDTA were added and the cells gently mixed for 20 minutes at room temperature. Following this 14ml of triton mix (0.1% triton X-100, 62.5mM EDTA, 50mM Tris.HCl pH8) was added and the samples incubated on ice with occasional mixing for 20 minutes. Cell debris was sedimented by centrifugation (39,000g, 45 minutes) allowing recovery of the supernate. Closed circular plasmid DNA was purified by CsCl equilibrium-gradient centrifugation as described by Maniatis *et al.*, 1982. CsCl (0.95g) and 0.1ml ethidium bromide (5mg/ml) were added to every 1ml of the supernate and the solution transferred into a 1.6 x 7.6cm polyallomer tube (Beckman). This was then sealed and centrifuged in a Sorvall OTD 50B ultracentrifuge using a Sorvall Ti50 rotor at 45,000rpm for at least 2 days at roughly 20°C . The tube was then viewed using UV light and the plasmid DNA (lower of the two bands) recovered through the side using a 21# gauge hypodermic needle and syringe. Ethidium bromide was removed by several extractions of the DNA with butanol (equilibrated with CsCl saturated TE). The CsCl was removed by dialysis against four changes of 2 litres of TE at 4°C The isolated DNA-containing solution was then ethanol precipitated, washed with 66% (v/v) ethanol and dried (vacuum desiccator). Following this, it was dissolved in TE, phenol/chloroform extracted and ethanol precipitated again. After washing the pelleted nucleic acid with 66% (v/v) ethanol, it was dried (vacuum desiccator) and re-dissolved in 500 $\mu$ l-1ml of TE.

#### 2.2.5.4 Preparation of M13 RF DNA

A single M13 plaque plus 5 $\mu$ l of a stationary *E.coli* culture (JM101 or NM522) was inoculated into 2.5ml of LB and incubated at 37°C with vigorous shaking for 6 hours. This culture was then stored overnight at 4°C The next day it was added to a freshly grown JM101

or NM522 culture (500ml; O.D. 550<sub>nm</sub> of 0.3-0.4) and incubated at 37 C for 4-4.5 hours. After this the protocol for isolation of plasmid DNA (see Section 2.2.5.3) was followed.

#### 2.2.5.5 Restriction Enzyme Digestion of DNA

DNA was incubated with the relevant enzyme(s) in a buffer and at a temperature indicated by the manufacturer. The duration of incubation, the number of enzyme units used and the total reaction volume was dependent on the amount of DNA to be restricted. Normally 1-5 $\mu$ g of DNA was restricted with 2-10 units of restriction enzyme, in 20 $\mu$ l for 1-4 hours. To avoid glycerol inhibition of enzyme digestion, the enzyme volume used never exceeded 0.1 volume of the total reaction.

#### 2.2.5.6 Agarose Gel Electrophoresis

In order to separate restricted DNA fragments and visualise nucleic acids, DNA and RNA were electrophoresed through horizontal agarose gels containing and submerged in TAE buffer (40mM Tris-acetate, 2mM EDTA). Linear DNA molecules in the size ranges of 7 to 0.5kb and 4 to 0.2kb were separated using 0.8% and 1.5% (w/v) gels respectively. Submerged 20 x 20cm gels were used for all large scale isolations of restricted DNA fragments whereas the 14 x 11cm gel kits supplied by BRL were utilised for all other purposes. To prevent light-induced nicking of DNA the large gels were electrophoresed at 10 V/cm in the absence of ethidium bromide and subsequently stained (0.1 $\mu$ g/ml ethidium bromide) in the dark. The small gels were electrophoresed at 5-10 V/cm in the presence of 0.1 $\mu$ g/ml ethidium bromide. The nucleic acid contained in these gels was observed using a short wave length UV transilluminator and photographed as indicated in Section 2.2.1.5.

#### 2.2.5.7 Isolation of Restricted DNA Fragments

Relevant restriction fragments were cut from agarose gels as thin slices and when

necessary covered with Saran wrap and tin-foil and stored at 4°C. The DNA was extracted from these slices in two ways;

(a) the Schleicher and Schuell Biotrap:-

The procedure followed is detailed in the booklet accompanying this apparatus. Briefly, the gel slice was placed in an electrophoresis chamber containing TAE and a current (10 V/cm) was passed through it. In the course of electrophoresis (2-4 hours) the DNA migrated into a collecting chamber and was there retained by binding specifically to a membrane. Reversal of the current for 30secs eluted the DNA from the retention membrane (without the exit of the DNA from the collecting chamber) and allowed its isolation in a relatively small volume (200-500µl). The DNA was then phenol/chloroform extracted and ethanol precipitated (see Sections 2.2.4.4 and 2.2.4.3). This procedure consistently produced >75% recovery of the isolated fragment (as judged by its intensity in ethidium stained gels).

(b) BIO101 Inc. GeneClean Kit

This commercial method is based on the ground glass procedure as described by Vogelstein and Gillespie (1979). The agarose gel matrix of the gel slice is dissolved by incubation in sodium iodide. DNA fragments in the resulting solution are free to adhere to an added silica matrix and can be sedimented out of the solution by centrifugation (17000g, 30secs). The DNA fragments can then be eluted from the silica into dH<sub>2</sub>O by heating the mixture at 45-55°C for a few minutes. A more detailed protocol is specified in the BIO101 Inc. GeneClean kit. This procedure is very fast (taking only 20 minutes) and produces efficient fragment recovery.

#### 2.2.5.8 Ligation of DNA

Linearised vector (final concentration of 10ng/µl) and insert DNA bearing complementary cohesive ends were mixed in a molar ratio of 1 : 1 in a total reaction volume of 20µl containing 4µl 5 X ligation buffer (330mM Tris.HCl pH7.5, 50mM MgCl<sub>2</sub>, 250mM



NaCl, 50mM DTT, 3mM rATP) and 5 units of T4 DNA ligase. This was incubated at 15°C for 12-16 hours and if necessary the ligase was then inactivated by incubation at 65°C for 10-15 minutes.

#### 2.2.5.9 Filling in DNA Termini With Recessed 3' Ends

Termini with recessed 3' ends were rendered "blunt" ended by incubation of 1-5µg of DNA in a 20µl reaction volume containing 2µl X 10 nick translation buffer (0.5M Tris.HCl. pH7.5, 0.1M MgCl<sub>2</sub>, 1mM DTT, 500µg/ml BSA), 0.1mM dNTPs and 2 units of the Klenow fragment of the *E.coli* DNA polymerase I at room temperature for 30 minutes. The reaction was stopped by incubation at 65°C for 10-15 minutes.

#### 2.2.5.10 Labelling DNA by the Random Priming Method

The random priming procedure was essentially performed as described by Feinberg and Vogelstein (1984). The reaction was assembled by the addition of each constituent in the stated order; dH<sub>2</sub>O to a final volume of 50µl, 10µl 5 X OLB (1M HEPES pH6.6, 25mM MgCl<sub>2</sub>, 50mM 2-mercaptoethanol, 250mM Tris.HCl. pH8, 0.1mM dNTPs, 7.5mg/ml random hexadeoxyribonucleotides), BSA to a final concentration of 0.4mg/ml, 50-100ng of heat-denatured DNA (see below), 3µl of [ $\alpha$ -<sup>32</sup>P]-dCTP (3000 Ci/mmol, 10µCi/µl); and 4 units of the Klenow fragment of the DNA polymerase I of *E.coli*. The reaction was incubated at room temperature for 5-6 hours and then the DNA was heat-denatured and added to the hybridisation reaction.

Gel-purified DNA fragments and radioactively labelled probe were heat-denatured by incubation at 100°C in a boiling water-bath for 5 minutes.

#### 2.2.5.11 Measurement of Radioactivity Incorporated into DNA

The efficiency of incorporation of radiolabel into DNA (see Section 2.2.5.10) was estimated by measuring acid-precipitable counts. A 1 $\mu$ l aliquot of the labelling reaction (see Section 2.2.5.10) was added to an eppendorf tube containing 45 $\mu$ l TE and 1 $\mu$ l 2mg/ml tRNA, after which 13 $\mu$ l of 50% (w/v) TCA was added and the mixture incubated on ice for 20-30 minutes. Precipitated DNA was collected by vacuum filtration onto a Whatman GF/C glass fibre disc (pre-wet with 10% (w/v) TCA), then washed with 10ml ice-cold 10% (w/v) TCA and 10ml ice-cold ethanol. As a control 1 $\mu$ l of the labelling reaction was spotted directly onto a GF/C disc. The amount of radioactivity present on each disc was determined by scintillation-counting and the percentage incorporation of radioactive label was estimated by comparing the two samples.

#### 2.2.5.12 5' End Labelling of Oligodeoxynucleotides

Oligodeoxynucleotides were radioactively labelled by the addition of a radioactive phosphate moiety to the 5' hydroxyl group present on these molecules. The reaction was carried out in a 30 $\mu$ l total volume containing 20 $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]rATP (3000 Ci/mmol, 10 $\mu$ Ci/ $\mu$ l), 0.1M Tris.HCl, pH8, 5mM DTT, 10mM MgCl<sub>2</sub>, 20pmol oligonucleotide and 5 units T4 polynucleotide kinase. The reaction was incubated at 37°C for 1 hour. The oligodeoxynucleotide is in excess in this reaction as the amount of [ $\gamma$ -<sup>32</sup>P]rATP used is roughly 6.7pmol (10 $\mu$ Ci = 3.35pmol rATP). The integrity of the labelled oligodeoxynucleotide was determined by PAGE. Briefly, 0.25 volume of sequencing load buffer (see Section 2.2.7.3) was added to an aliquot of the labelled oligodeoxynucleotide, boiled for 5 minutes, loaded onto a 20% polyacrylamide gel (see below) and electrophoresed at 33W for 2-4 hours. One of the glass plates was removed and the gel and other glass plate covered in Saran wrap and exposed to X-ray film (20% gels cannot be transferred to filter paper). In this way the size of the oligodeoxynucleotide and its ability to be phosphorylated were determined.

The polyacrylamide gel was made thus: a 40% stock acrylamide solution was made by adding 10g of bis-acrylamide to 190g of acrylamide. Urea (240g) was then added (final concentration 7.8M) and the constituents were dissolved in dH<sub>2</sub>O to a final volume of 500ml. This solution was deionised (see Section 2.2.1.3) and stored at 4°C. For a 20% gel, 25ml of this solution was added to 20ml of an 8M urea solution and 5ml of 10 X TBE. Polymerisation and pouring of such gels is detailed in Section 2.2.7.4.

#### 2.2.5.13 Purification of Labelled Oligodeoxynucleotide

When necessary the labelled oligodeoxynucleotide was separated from the unincorporated label by gel exclusion chromatography. The labelled oligodeoxynucleotide (see Section 2.2.5.12) was brought to a volume of 1ml with elution buffer (10mM sodium phosphate pH7), added to a pre-equilibrated Sephadex G-25 NAP-10 column (Pharmacia ;column volume 1.5mls) and eluted in the second 1.5mls aliquot washed through the column.

#### 2.2.5.14 Colony Blotting and Hybridisation

Colony screening was essentially performed as described by Maniatis *et al.*, (1982). Transformed *E.coli* colonies were arranged in a grid pattern on LB-amp plates and incubated overnight at 37°C. A Hybond-N nylon membrane (Amersham) was placed on top of the colonies, marked to indicate its orientation and left at room temperature for 2-3 minutes. The membrane was then removed and placed (colony side up) onto filter paper soaked in Southern Denaturation buffer (1.5M NaCl, 0.5M NaOH) for three minutes, then onto fresh paper soaked in the same buffer for a further three minutes. Excess buffer was then removed by blotting on filter paper and the membrane soaked twice (colony side up; 5 minute durations) on filter paper saturated with Neutralisation buffer (0.5M Tris.HCl pH7.5, 1.5M NaCl). The membrane was

"TE" buffer : 10mM Tris.HCl pH8  
1mM EDTA

"TE"<sub>0.1</sub> buffer : 10mM Tris.HCl pH8  
0.1mM EDTA

5 X "P" buffer : 1% (w/v) BSA  
1% (w/v) Polyvinylpyrrolidone (PVP)  
1% (w/v) Ficoll  
250mM Tris.HCl pH7.5  
0.5% sodium pyrophosphate  
5% SDS

5 X Denhardt's : 0.1% (w/v) BSA  
solution 0.1% (w/v) PVP  
0.1% (w/v) Ficoll

---

then briefly washed in 2 X SSC, dried at room temperature, covered in Saran wrap and the DNA cross-linked to the membrane by placing it (colony side down) on a short-wave length UV transilluminator for 5 minutes. The membrane was then placed in a plastic bag containing 10-25mls of pre-hybridisation solution (2 X SSC, 10 X Denhardt's solution, 0.1mg/ml heat-denatured salmon sperm DNA), sealed, placed in another bag (which was then sealed) and incubated at 65°C (with constant agitation) for 3-4 hours. The pre-hybridisation solution was then poured off and replaced with 10-25mls hybridisation solution (pre-hybridisation solution minus the salmon sperm DNA). Heat-denatured probe was added to the bag, which was re-sealed and incubated at 65°C with constant agitation for 12-16 hours. After this the membrane was recovered from the bag and non-specifically bound probe removed by two sequential 10 minute washes in 500ml of 2 X SSC at room temperature, followed by two 30 minute washes in 500ml of 2 X SSC, 0.5% SDS at 65°C. If required a further two 10 minute washes in 500ml of 0.1% SSC at room temperature were performed. Excess SSC was removed by blotting the membrane on filter paper, which was then covered with Saran wrap and exposed to X-ray film (see Section 2.2.1.4).

#### 2.2.5.15 M13 Plaque Hybridisation Using DNA Fragment Probes

The procedure followed was as described by Mason and Williams, (1985). BBL top agar plates harbouring M13 plaques were incubated at 4°C for >30 minutes which helped to reduce the amount of top agar that stuck to the membrane when the plaques were lifted. A nitrocellulose membrane was placed over the plaques, marked and orientated, and left at room temperature for 5 minutes to allow adsorption of the plaques to the membrane. The membrane was then baked at 80°C in a vacuum oven for 2-4 hours. No denaturation/neutralisation steps were required as M13 contains a single-stranded DNA genome; presumably the baking step denatures the protein coat to enable the probe to hybridise to the DNA. After baking, the membrane was briefly soaked in SET buffer (3M NaCl, 0.4M Tris.HCl pH7.5, 20mM EDTA) making sure both sides of the membrane were wet, blotted onto filter paper to remove excess

"TE" buffer : 10mM Tris.HCl pH8  
1mM EDTA

"TE"<sub>0.1</sub> buffer : 10mM Tris.HCl pH8  
0.1mM EDTA

5 X "P" buffer : 1% (w/v) BSA  
1% (w/v) Polyvinylpyrrolidone (PVP)  
1% (w/v) Ficoll  
250mM Tris.HCl pH7.5  
0.5% sodium pyrophosphate  
5% SDS

5 X Denhardt's : 0.1% (w/v) BSA  
solution 0.1% (w/v) PVP  
0.1% (w/v) Ficoll

---

buffer and sealed in a plastic bag containing 10mls of pre-hybridisation solution (5ml formamide, 2ml 5 X "P" buffer, 2.5ml 4M NaCl, 0.1ml 10mg/ml heat-denatured salmon sperm DNA). This was incubated with constant agitation at 42°C for 2-3 hours. The pre-hybridisation solution was replaced with 10ml hybridisation solution (pre-hybridisation solution minus the salmon sperm), the probe (heat-denatured) was added and the bag was incubated at 42°C with constant agitation for 12-16 hours. It should be noted that during the pre- and hybridisation steps the membrane was sealed within two bags (minimum). Removal of non-specifically bound probe and exposure to X-ray film was as described in Section 2.2.5.14.

#### 2.2.5.16 Plaque Hybridisation Using Oligodeoxynucleotide Probes

The procedure followed was a modification of that described by Zoller and Smith, 1984 and relates only to site-directed mutagenesis. M13 plaques adsorbed to a Hybond-N nylon membrane (see Section 2.2.5.15) were then placed (plaque side up) on filter paper soaked in Southern Denaturation buffer (1.5M NaCl, 0.5M NaOH) for 7 minutes. This process was repeated on freshly soaked paper. Excess buffer was removed by blotting on filter paper, then the membrane was twice soaked on filter paper saturated with Neutralisation buffer (0.5M Tris.HCl pH7.5, 1.5M NaCl: plaque side up; 3 minute durations) and briefly washed in 2 X SSC. After drying (37°C, 15 minutes), the membrane was covered with Saran wrap and the DNA cross-linked to it by exposure to short wave-length UV light (plaque side down) for 5 minutes. The membrane was then placed in a plastic tray containing 25ml of pre-hybridisation solution (6 X SSC, 5 X Denhardt's solution, 0.5% SDS, 0.1mg/ml heat-denatured salmon sperm DNA) and incubated with constant agitation at 65°C for at least 1 hour. If more than one membrane was pre-hybridised then 25ml of pre-hybridisation solution/membrane was added. The membrane was then placed in a plastic bag and 10ml of hybridisation solution added (6 X SSC, 5 X Denhardt's solution). The labelled oligodeoxynucleotide was separated from unincorporated label as described in Section 2.2.5.13 and added straight to the bag containing the membrane and hybridisation solution. If more than one membrane was used they were placed in separate

bags and each received 20pmol of radioactively labelled probe (see Section 2.2.5.12) and 10ml hybridisation solution. The membrane was then incubated with constant agitation at room temperature for at least 4 hours. After recovery the membrane was washed three times in 100ml of 6 X SSC at room temperature (3 minute durations) and excess buffer was removed by blotting on filter paper. The membrane was then covered with Saran wrap and exposed to X-ray film. Subsequent washes were carried out in 6 X SSC at increasing temperature (see Chapter 3.3.3.1).

It should be noted that after pre-hybridisation the membrane was never allowed to dry completely; blotting on filter paper was used only as a means to remove excess liquid. Dry membranes bind oligonucleotides irreversibly and make it impossible to differentially remove mismatched hybrids by increasing the wash temperature (see Chapter 3.3.3.1 for a fuller description of the wash process).

## 2.2.6 Site-Directed Mutagenesis

### 2.2.6.1 5' Phosphorylation of Mutagenic Oligodeoxynucleotides

Stock solutions of phosphorylated oligodeoxynucleotides were prepared in the following manner: in a reaction volume of 30 $\mu$ l, 200pmol of oligonucleotide was added to 0.1M Tris.HCl pH8, 10mM MgCl<sub>2</sub>, 5mM DTT, 1mM rATP and 5 units T4 polynucleotide kinase. This was incubated at 37°C for 1 hour. If required the T4 polynucleotide kinase was inactivated by incubation at 65°C for 10-15 minutes.

### 2.2.6.2 Annealing Conditions

Efficient recovery of M13 phage harbouring site-directed mutations relies on the specific hybridisation of the mutagenic oligodeoxynucleotide to the DNA sequence to be mutated. The number of mutants recovered will increase with the number of oligonucleotide



molecules that hybridise to the sequence of interest. However if you present a vast excess of oligodeoxynucleotide molecules (relative to the sequence of interest), non-specific hybridisation will reduce the efficiency of mutant recovery. Another reason for such a reduction is non-specific hybridisation which is a result of complementarity between the mutagenic oligodeoxynucleotide and other sequences distinct from the one to be mutated. It is therefore important to obtain hybridisation that is specific to the sequence to be mutated. Non-specific hybridisation can be reduced not only by making sure that the oligodeoxynucleotide : template ratio is correct but also by increasing the stringency of the annealing reaction (i.e. by increasing the temperature); as the annealing temperature increases only those molecules with the highest degree of complementarity will remain hybridised to their complementary sequence. Therefore both the ratio of oligodeoxynucleotide molecules to template molecules and the temperature at which they are annealed are controllable variables that can alter the pattern of hybridisation.

By using the mutagenic oligodeoxynucleotides as sequencing primers it was possible to determine whether they annealed at the correct place. Correctly annealed oligodeoxynucleotides produced the same sequencing pattern as the M13 universal primer: of course the M13 primer initiated polymerisation 5' to the oligodeoxynucleotide and therefore contained sequences not detected by the oligodeoxynucleotide. Oligodeoxynucleotides annealing at various regions of the template DNA produced smeared sequencing patterns, representing the multiple sizes of terminated fragments. In this way the temperature of annealing and the oligodeoxynucleotide : template ratio were assayed.

Normal sequencing reactions were performed (see Section 2.2.7.3) except that only the (A) sequencing solution was used (i.e. A-tracking), therefore ingredients in the method were reduced accordingly. Template DNA (0.125pmol; 1 $\mu$ l of normal template preparation) was mixed with 10 X, 25 X and 50 X excess of oligodeoxynucleotide (1 $\mu$ g solution containing 1.25, 3.125 and 6.25pmol respectively and the other ingredients of the annealing mix (AM) see Section 2.2.7.2) and each was incubated at 55°C, 70°C and 85°C for 15 minutes. The samples were then briefly spun (5 secs; 17,500g) and incubated at room temperature for 15 minutes. After this annealing reaction, 2.9 $\mu$ l of a solution containing 0.4 $\mu$ l of [ $\alpha$ -<sup>35</sup>S]-dATP (400 Ci/mmol) and

2.5 $\mu$ l of sequencing solution (A) was added and mixed. To this 0.5 units of DNA polymerase I of *E.coli* was added (Klenow fragment, in a volume of 2 $\mu$ l made by making a 1 in 4 dilution of stock Klenow in cold 10mM Tris.HCl pH7.5). The samples were mixed and incubated at room temperature for 20 minutes, followed by the addition of 2.5 $\mu$ l of Chase Mix (1mM dNTPs) and further incubation at room temperature for 20 minutes. Sequencing load buffer (4 $\mu$ l) was then added and the samples were incubated at 100°C for 5 minutes and fractionated by polyacrylamide gel electrophoresis (see Section 2.2.7.4). Fixing of the gel and its subsequent exposure to X-ray film were as described in Section 2.2.7.4.

#### 2.2.6.3 Site-Directed Mutagenesis: Annealing and Extension

In a total volume of 10 $\mu$ l, template DNA (0.5pmol; 4 $\mu$ l of normal template preparation) was incubated with the mutagenic oligodeoxynucleotide and 1 $\mu$ l 10 X buffer A (0.2M Tris.HCl pH7.5, 0.1M MgCl<sub>2</sub>, 0.5M NaCl, 10mM DTT) at a temperature and ratio previously determined (see Section 2.2.6.2). Annealing continued at the relevant temperature for 15 minutes and the samples were then centrifuged (17,500g, 10secs), mixed and incubated at room temperature for another 15 minutes. After this annealing reaction, 10 $\mu$ l of E/L mix (1 $\mu$ l 10 X buffer B {0.2M Tris.HCl pH7.5, 0.1M MgCl<sub>2</sub>, 0.1M DTT}, 0.8mM dNTPs, 1mM rATP, 3 units T4 ligase, 2 units Klenow) was added and the mixture was incubated at 16°C for 12-16 hours.

#### 2.2.6.4 Transformation

An aliquot of the E/L mix (5 $\mu$ l) was transformed into the *E.coli* strain BMH71-18 *mutL* (see Section 2.1.3) and 2-5 $\mu$ l aliquots of the transformation mix were plated along with 200 $\mu$ l of JM101 plating cells, 30 $\mu$ l X-gal (20mg/ml), 20 $\mu$ l IPTG (24mg/ml) and 3.5ml BBL top agar (see Section 2.2.2.3). The BMH71-18 *mutL* strain is defective in mismatch repair (Kramer *et al.*, 1984) and therefore the DNA heteroduplex introduced i.e. the duplex

containing the mismatch, is unlikely to be corrected to the wild-type sequence. Thus the specific mutation is "fixed" in BMH71-18 *mutL*. However, it is possible that spontaneous mutations may arise in other regions and remain uncorrected. M13 plaques arise by continuous infection of *E.coli* cells and if an excess of JM101 plating cells (repair proficient) is added it is likely that they will be infected in preference to the relatively small amount of BMH71-18 *mutL* cells present. As a consequence, spontaneous mutations that arise later in the plaque life-cycle (i.e after the first round of infection) probably do so in JM101 and therefore have more chance of being corrected to wild-type. It was hoped that the frequency of spontaneous mutations would be minimised using this approach.

Detection of plaques containing mutant phage is described in Chapter 3.3.3.1 and Section 2.2.5.16.

## 2.2.7 DNA Sequencing

### 2.2.7.1 Preparation of Template DNA

A 1% (v/v) inoculum from a fresh stationary culture of *E.coli* strains JM101 or NM522 was grown at 37°C in LB until an OD<sub>550nm</sub> of 0.2-0.3 was obtained. The cells were then aliquoted (1.5ml) into 5ml glass bottles, to which single M13 plaques were added. The bottles were incubated at 37°C with vigorous shaking for 4-4.5 hours and then the cultures transferred to eppendorf tubes. The cells were sedimented by centrifugation (2 x 17,500g, 5 minutes each time), the supernate was recovered and 150µl of a solution containing 2.5M NaCl and 20% (w/v) polyethylene glycol MW 6,000 (PEG) was added. After a 20 minute incubation at room temperature, precipitated phage were pelleted by centrifugation (17,500g, 10 minutes) and the supernate discarded. The tubes were then briefly respun and all traces of liquid removed. The phage pellet was then dissolved in 100µl TE0.1, to which 50µl of phenol (equilibrated with TE0.1) was added. The samples were then vortexed (10 secs), incubated at room temperature (5 minutes), re-vortexed (10 secs) and then the phases

separated by centrifugation (17,500g, 5 minutes). After recovery of the aqueous phase, 10 $\mu$ l of 3M sodium acetate and 250 $\mu$ l of ethanol were added, followed by incubation at -20°C for 15 minutes. Precipitated DNA was then pelleted by centrifugation (17,500g, 30 minutes), washed with 66% (v/v) ethanol, re-spun (17,500g, 10 minutes), dried (vacuum desiccator) and resuspended in 25 $\mu$ l TE0.1. The concentration of DNA in the final sample was roughly 0.125pmol/ $\mu$ l.

#### 2.2.7.2 Sequencing Solutions

Annealing Mix (AM) was made by mixing heptadecamer sequencing primer (2.5 $\mu$ g, New England Biolabs), 10 X TM buffer (0.1M Tris.HCl pH7.5, 0.1M MgCl<sub>2</sub>) and dH<sub>2</sub>O in the ratio 2:3:7.

Sequencing solutions G, C and T were made by mixing dG, dC and dT (see Table 2.5) with an equal volume of ddG, ddC and ddT (see Table 2.6) respectively. Sequencing solution A was made by mixing two volumes of dA with one volume of ddA. All solutions were stored at -70 C.

#### 2.2.7.3 Sequencing Reactions

Template DNA (0.5pmol = 4 $\mu$ l of template preparation) was mixed with 4 $\mu$ l of AM in an eppendorf tube, the cap was tightly sealed and then the tube placed in a water bath pre-heated to 70°C. The samples were allowed to cool in the water-bath for 15 minutes, centrifuged briefly (17,500g, 10 secs), mixed again and incubated at room temperature for a further 15 minutes. During these incubation periods, 2.5 $\mu$ l of the sequencing solutions A, G, C and T were aliquoted into 4 eppendorf tubes marked "A", "G", "C" and "T" respectively. After incubation, 1.5 $\mu$ l of [ $\alpha$ -35S]-dATP (400Ci/mmol) and 2 units of *E.coli* DNA polymerase I (Klenow fragment) were mixed with the annealed template. Aliquots of this mixture (2.5 $\mu$ l) were then allocated to the 4 separate tubes containing either A, G, C

or T sequencing solutions, mixed and incubated at room temperature for 20 minutes. Chase Mix (1mM dNTPs) was then added (2.5 $\mu$ l), followed by incubation at room temperature for 20 minutes. After this, 4 $\mu$ l of sequencing load buffer (3ml deionised formamide, 300 $\mu$ l of 3% (w/v) xylene cyanol, 300 $\mu$ l of 3% (w/v) bromophenol blue, 12 $\mu$ l of 0.5M EDTA) was added and the samples were incubated at 100°C for 5 minutes and then fractionated on a 6% polyacrylamide gel (see Section 2.2.7.4

TABLE 2.5 COMPOSITION OF dNTP SOLUTIONS

(all volumes in  $\mu\text{l}$ )dNTP Solutions

<u>Stock Solutions</u>	<u>dA</u>	<u>dG</u>	<u>dC</u>	<u>dT</u>
10mM dGTP	10	-	10	10
10mM dCTP	10	10	-	10
10mM dTTP	10	10	10	-
0.5mM dGTP	-	10	-	-
0.5mM dCTP	-	-	10	-
0.5mM dTTP	-	-	-	10
TE0.1	570	570	570	570

TABLE 2.6 COMPOSITION OF ddNTP SOLUTIONS

(all volumes in  $\mu\text{l}$ )ddNTP Solutions

<u>Stock Solutions</u>	<u>ddA</u>	<u>ddG</u>	<u>ddC</u>	<u>ddT</u>
10mM ddATP	2	-	-	-
10mM ddGTP	-	6	-	-
10mM ddCTP	-	-	2	-
10mM ddTTP	-	-	-	10
TE0.1	198	198	194	190

#### 2.2.7.4 Sequencing Gels

Sequencing gels (6% (w/v) polyacrylamide, 7.8M urea) were prepared as follows: 3g bis-acrylamide, 57g acrylamide and 470g of urea were dissolved in 600ml of dH<sub>2</sub>O. The volume was brought to 900ml with dH<sub>2</sub>O, deionised and stored at 4°C. When required 45ml of this stock plus 5ml 10 X TBE (0.89M Tris.borate, 0.89M boric acid, 20mM EDTA) was added to a 50ml conical flask, to which 400µl 10% (w/v) APS and 40µl TEMED were then added. The contents were mixed and poured between two (42 X 23cm) glass plates which were separated along their 42cm edges by 0.5 X 42cm strips of 0.4mm modellers plasticard, and sealed along three edges with PVC tape. A shark-tooth comb was inserted (straight side placed into the polyacrylamide) in between the two plates and polymerisation of the gel was allowed to proceed overnight. After this the sealing tape along the bottom of the plates and the shark-tooth comb were removed and the comb replaced with its teeth touching the the top of the gel. This was then assembled into a vertical electrophoresis apparatus containing 1 X TBE. Prior to loading the gel was electrophoresed at 33 Watts for 30 minutes to warm it to its operational temperature. The sample-wells were flushed out with TBE using a pasteur pipette and heat-denatured samples were then loaded using a drawn-out capillary tube. Electrophoresis was performed at 33 watts (27-48 V/cm) for 2-4 hours. Following this the gel plates were separated and the plate attached to the gel was placed horizontally into two litres of "fix" (10% (v/v) acetic acid, 10% (v/v) methanol) for 15 minutes. After removing the plate from the "fix" the gel was transferred to blotting paper, covered with Saran wrap and dried in a vacuum gel drier at 80°C for 2 hours. The Saran wrap was then removed and the gel exposed to X-ray film (see Section 2.2.1.4).



## 2.2.8 RNA Methods

### 2.2.8.1 Protection of RNA from RNases

To preclude RNase contamination of RNA samples, all glassware was baked at 250°C overnight and all materials were manipulated with gloved hands. Solutions used in experiments involving RNA were rendered RNase-free by mixing with 0.1% (v/v) diethylpyrocarbonate (DEP) for 1-2 hours at room temperature, followed by autoclaving (see Section 2.2.1.2). Solutions containing EDTA or Tris were not treated with DEP because it reacts chemically with the amino groups of these compounds. Instead, these solutions were prepared in RNase-free glassware with DEP treated, autoclaved dH<sub>2</sub>O.

### 2.2.8.2 Extraction of RNA from Yeast

Yeast cultures (200ml) were grown at 23°C until the OD 600<sub>nm</sub> was between 0.2 and 0.7. At this stage (if required) the culture was split and one half heat-shocked thus: an equal volume of pre-incubated medium (49°C) was added (final temperature, 36°C) and incubated with vigorous shaking at 36°C for the required time. The other half of the culture was treated in a similar manner except that the medium was pre-incubated at 23°C (final temperature, 23°C).

All of the following steps were performed at 4°C or on ice. The cultures were then added to an equal volume of ice/water, sedimented by centrifugation (4,100g, 7 minutes), resuspended in 9ml ice-cold TNE (50mM Tris.HCl pH7.5, 100mM NaCl, 5mM EDTA) and transferred to a 15ml Corex tube. The cells were sedimented by centrifugation (4,100g, 7 minutes), resuspended in 1ml TNE and 30mg of acid-washed glass beads were added (see below). The cells were fractured by vortexing vigorously for 1 minute; 4ml TNE, 0.2ml 20% SDS and 4ml phenol (equilibrated with DEP treated dH<sub>2</sub>O) were added rapidly, followed by further vortexing for 1.5 minutes. After centrifugation (4,100g, 5 minutes)

the aqueous phase was recovered, extracted with phenol/chloroform (1:1) and chloroform, and then the RNA was precipitated with ethanol (see Section 2.2.4.3). The integrity of the RNA was determined by electrophoresing samples in agarose gels containing ethidium bromide (see Section 2.2.5.6).

Acid-washed glass beads were prepared by heating 10g glass beads (approximately 100 mesh, BDH) with 100ml 1M HCl to 100°C for 20 minutes. The beads were then washed sequentially with 1 litre dH<sub>2</sub>O, 300ml of both 1M and 0.1M Tris.HCl pH7.5 and then 2 litres of dH<sub>2</sub>O. They were then dried at 65°C (overnight) and baked at 250°C (overnight).

### 2.2.8.3 RNA Agarose Gel Electrophoresis

Horizontal slab gel electrophoresis of RNA was performed in submerged 20 X 20cm, 1.5% (w/v) agarose gels containing 10mM sodium phosphate pH7 (the electrophoresis buffer was also 10mM sodium phosphate). Before loading, the RNA samples and the DNA markers were denatured by incubation at 50°C for 1 hour with 4 volumes glyoxal mix (25% (v/v) glyoxal, 62.5% (v/v) deionised formamide, 12.5mM sodium phosphate pH7). This procedure is a modification of that described by McMaster and Carmichael, (1977). Samples were electrophoresed at 10V/cm.

### 2.2.8.4 Northern Blotting and Hybridisation

Following electrophoretic separation of RNA, the gel was assembled into a capillary blot. A glass plate was placed over a reservoir containing approximately 1 litre of 1 X SSC. Two sheets of Whatman 3MM filter paper (pre-soaked in 1 X SSC) were then placed on top of the plate with their ends submerged in the reservoir. The gel was placed on top of the paper (sample-wells facing down) and PVC tape placed all around to inhibit buffer transfer that would bypass the gel and membrane. On top of this were placed sequentially; pre-soaked membrane (Genescreen, see below), 3 double sheets of Whatman 3MM paper

(soaked in 1 X SSC), 3 double sheets of Whatman 3MM paper (dry), a 10-15cm stack of paper towels and a 500g weight. During blot assembly, care was taken to remove air bubbles present between the wetted paper, the gel and membrane, and the membrane/paper interfaces. After leaving for roughly 16 hours the membrane was removed, washed briefly in 1 X SSC, dried at 37°C for 15 minutes and then baked at 80°C for 2 hours in a vacuum oven. Pre-hybridisation and hybridisation were performed as described in Section 2.2.5.15, except that a Hybaid hybridisation oven, employing bottles instead of plastic bags, was used (see Hybaid hybridisation oven manual for instructions). Washes and exposure to X-ray film were as described in Section 2.2.5.14.

Prior to being placed in the blot, Genescreen membranes were soaked in dH<sub>2</sub>O (2 minutes) and 1 X SSC (30 minutes).

#### 2.2.8.5 Primer Extension

The procedure followed was a modification of that described by Domdey *et al.*, (1984) and also detailed in the GIBCO BRL catalogue. End-labelling (5') of the oligodeoxynucleotide was as described in Section 2.2.5.12, except that 7pmol of oligodeoxynucleotide was used. The radioactively labelled oligodeoxynucleotide (0.667pmol) was annealed to 30-100µg total RNA, at a temperature equal to  $T_m - 5^{\circ}\text{C}$  for 15-30 minutes. This reaction was carried out in a 30µl total volume containing 50mM Tris.HCl pH7.5, 75mM KCl and 3mM MgCl<sub>2</sub>. The samples were then incubated on ice for 5 minutes after which 19µl of a solution containing 50mM Tris.HCl pH7.5, 75mM KCl, 5mM MgCl<sub>2</sub>, 25mM DTT, 1.25mM dNTPs and 0.25mg/ml BSA (high grade) was added. Following this, 100-200 units (1-2µl) of Moloney Murine Leukemia Virus Reverse Transcriptase were added and the samples were mixed and incubated at 37°C for 1 hour. To each sample 12.5µl of 0.5M NaOH were then added, followed by incubation at 100°C for 3 minutes and immediate incubation on ice (5 minutes). After briefly centrifuging the samples (17,500g, 10 secs), 12.5µl of 0.5M HCl and 12.5µl of Tris.HCl pH7.5 were added and

the DNA was ethanol precipitated (see Section 2.2.4.3). After centrifugation (17,500g, 30 minutes) the DNA pellet was washed with 66% (v/v) ethanol, dried (vacuum desiccator) and suspended in 3 $\mu$ l dH<sub>2</sub>O/3 $\mu$ l 2 X urea load buffer (10M urea, 1 X TBE, 0.2% xylene cyanol and bromophenol blue). Prior to electrophoresis through a 6% (w/v) polyacrylamide gel (see Section 2.2.7.4), the samples were incubated at 100°C for 4 minutes and briefly centrifuged (17,500g, 10secs). Following electrophoresis the glass plates were removed, the gel was transferred to filter paper (no gel "fixing" is required for nucleic acid labelled with [<sup>32</sup>P]; see Section 2.2.7.4), covered with Saran wrap and exposed to X-ray film (see Section 2.2.1.4).

## 2.2.9 $\beta$ -Galactosidase Assays

### 2.2.9.1 Induction of the *GAL1* Promoter

Yeast strains harbouring plasmids with the *GAL1* promoter, were grown in YMR/G/Cas at 23°C until an OD 600<sub>nm</sub> of 0.1-0.4 was attained. The cultures were then split and one half heat-shocked thus: an equal volume of pre-incubated YMR/Cas (49°C) was added (final temperature 36°C) and the culture incubated with vigorous shaking at 36°C. The other half of the culture was treated in a similar manner except that the YMR/Cas was pre-incubated at 23°C (final temperature 23°C). If induction of the promoter was to occur at 30°C then the cultures were grown at 30°C and the YMR/Cas pre-incubated at 30°C. Fifteen minutes after the addition of YMR/Cas, galactose (to a final concentration of 2%) was added and this point was regarded as time-zero. Thereafter aliquots of the cultures were removed at hourly intervals and assayed for  $\beta$ -galactosidase activity.

Induction of cultures from which RNA was extracted was performed thus: yeast cells were grown in 6 x 100ml of YMR/G/Cas at 23°C to an OD 600<sub>nm</sub> of 0.1-0.4. The cultures were not split, and 6 x 100ml of pre-incubated YMR/Cas (49°C or 23°C) was added. The

cultures were then incubated at the relevant temperature (36°C or 23°C). Fifteen minutes after the addition of YMR/Cas, galactose (2% final concentration; time-point zero) was added and thereafter one of the 6 cultures was removed at hourly intervals. The protocol for RNA preparation was followed prior to the cells being resuspended in 1ml TNE (see Section 2.2.8.2); at this stage they were resuspended in 2ml ethanol (-20°C) and stored at -20°C overnight (for at least 4 days RNA did not degrade under these conditions). The next day the cells were sedimented by centrifugation (4,100g, 5 minutes, 4°C), washed with 2ml ice-cold TNE, re-sedimented (4,100g, 5 minutes, 4°C) and resuspended in 1ml TNE. The normal protocol for RNA preparation was then continued.

#### 2.2.9.2 $\beta$ -Galactosidase Assays

The procedure used was a modification of that described by Legrain and Rosbash, 1989. Culture samples (0.25-2ml) to be assayed (see Section 2.2.9.1) were immediately sedimented by centrifugation (4,100g, 7 minutes), suspended in 200 $\mu$ l 50mM potassium phosphate pH7 and incubated at 28°C for 5 minutes. After this 200 $\mu$ l Z buffer (60mM di-sodium hydrogen orthophosphate, 40mM sodium di-hydrogen orthophosphate, 10mM KCl, 1mM magnesium sulphate, 50mM  $\beta$ -mercaptoethanol, 0.025% (w/v) SDS: pH7) and 2.5 $\mu$ l chloroform were added, then the samples were vortexed (10secs) and re-incubated at 28°C for 5 minutes. Following this, 200 $\mu$ l of 4mg/ml ONPG (dissolved in 0.1M potassium phosphate pH7) was added and the samples were mixed and then incubated at 28°C. The reaction was stopped by addition of 250 $\mu$ l of 1M sodium carbonate, the cell debris sedimented by centrifugation (4,100g, 7 minutes) and the  $\beta$ -galactosidase activity quantified by measuring the OD 420<sub>nm</sub> of the supernate. The units of  $\beta$ -galactosidase activity were calculated using the formula given by Miller, 1972 (see below).

$$\text{Units} = \frac{\text{OD } 420_{\text{nm}}}{t \times \text{OD } 600_{\text{nm}} \times v} \times 1000$$

OD 420<sub>nm</sub>      reflects the hydrolysis of ONPG by  $\beta$ -galactosidase  
OD 600<sub>nm</sub>      reflects the cell density of the culture just before the assay  
t                is the time of the reaction in minutes  
v                is the volume of culture used in the assay, in ml

This formula is not a measure of the specific activity of  $\beta$ -galactosidase, which is normally defined as  $\beta$ -galactosidase units/mg of protein. However, these units are proportional to the increase in o-nitrophenol per minute per yeast cell and, as it is assumed that the protein content of an exponentially growing culture is proportional to its cell density (i.e. its cell number), it is adequate for the purposes for which it was used. It should be noted that the spectrophotometer used was calibrated for both OD 420<sub>nm</sub> and 600<sub>nm</sub>, and all measurements were made within the linear range of the machine's capabilities.

## CHAPTER 3

# CONSTRUCTION OF PLASMIDS USED IN MUTANT SCREENING AND ANALYSIS OF PRE-mRNA COMMITMENT TO SPLICING

### 3.1 INTRODUCTION

This chapter describes the rationale for, and construction of, those plasmids required in setting up a *prp* mutant screen. In the course of these constructions it became apparent that two further plasmid constructs would allow a study of the commitment of pre-mRNA to splicing. This chapter describes the construction of all the plasmids, as they were derived from the same "parent" clone and made by a similar method. However the rationale for the construction of those plasmids involved in the study of pre-mRNA commitment is dealt with in the Appendix.

### 3.2 RATIONALE FOR CONSTRUCTION OF THE SCREENING PLASMIDS

The primary objective of this thesis was to develop a novel method of screening *t<sup>s</sup>* yeast mutants for defects in pre-mRNA processing at the restrictive temperature. The vast majority of *prp* mutants accumulate pre-mRNA with concomitant loss of mRNA and this is probably indicative of defects in splicing (for further discussion see Section 1.9). Large scale screening of *t<sup>s</sup>* mutants for such defects has relied exclusively on Northern blot analysis to detect pre-mRNA accumulation. *Prp* mutants are thought to account for approximately 5% of all *t<sup>s</sup>* mutants (Hartwell *et al.*, 1970; Vijayraghavan *et al.*, 1989). If a bank of *prp* mutants is desired, then such an approach is both time consuming and labour intensive, requiring RNA preparations from many *t<sup>s</sup>* mutants grown at the non-permissive and permissive temperatures (Vijayraghavan *et al.*, 1989; Warner, 1989; Hartwell, 1967). At the time of establishing a new method the only *prp* mutants in existence were *prp2-prp11*, which all accumulate pre-mRNA. It

was considered that any screen should rely on this common property, and if it was to be fast it had to obviate the need for extensive Northern blot analysis. The basic aim was therefore to measure pre-mRNA accumulation by another procedure. The approach taken was to try to detect any increase in pre-mRNA indirectly, by measurement of an assayable protein which was encoded by the pre-mRNA only. Exploiting the appropriate protein, a simple plate assay was to be developed to allow large scale screening. This screen therefore relied on three important criteria :

- (a) that only the pre-mRNA could encode the assayable protein
- (b) that pre-mRNA would be transported from the nucleus to the cytoplasm where it would be translated
- (c) that pre-mRNA accumulation would result in an increase in the assayable protein activity.

The rest of this chapter is concerned with the creation of suitable gene constructs that either accomodate criterion (a) above or are variations on a similar theme. The remaining points are addressed in Chapter 4 and the Appendix.

There is no known yeast gene in which only the pre-mRNA can encode an assayable protein (criterion (a) above), therefore direct assay of mutagenised cells was not possible. In order to use the approach outlined above a novel gene was required i.e. one that had an ORF through its pre-mRNA that encodes assayable activity, yet when spliced had none (i.e. the screening gene). Such a gene was to be introduced into yeast using a suitable plasmid vector (i.e. the screening construct).

This novel gene was created by fusing part of a yeast intron-containing gene to the 5' end of the *lacZ* gene of *E.coli*, thereby encoding a potential  $\beta$ -galactosidase fusion protein which initiates translation in exon 1 of the yeast gene. As it was a normal yeast gene it had the advantage that it possessed all the essential sequence elements required for accurate splicing, and the use of *lacZ* allowed both the biochemical analysis of cell extracts and the development of a colourmetric colony assay. The *MATa1* gene of yeast was utilised because it contained a



small IVS-1 (54 nucleotides) rendering it easily amenable to genetic manipulation (cloning and SDM), and its transcript was inefficiently spliced and therefore more sensitive to subtle changes in splicing efficiency. In this way it was hoped to isolate splicing mutants with relatively weak phenotypes. More importantly it had an ORF occurring through E1-IVS1-E2 allowing expression of a fusion protein when fused in-frame with *lacZ*. However to satisfy the criterion required of the screening gene (criterion (a) above), the intron had first to be modified such that, following splicing, the now intron-less *MATa1* region was out-of-frame with the *lacZ* ORF and could not encode an active  $\beta$ -galactosidase fusion protein.

To confirm that unspliced transcript encoded the active fusion protein a stop codon was engineered into the ORF of the intron. The resulting construct was called pJBM-4/s. A  $\beta$ -galactosidase fusion protein should not be encoded by the *MATa1-lacZ* gene fusion present in pJBM-4/s.

In screening *t<sup>s</sup>* mutants for a *prp* phenotype it was envisaged that  $\beta$ -galactosidase activity would be measured from yeast strains incubated at the permissive and non-permissive temperatures. Thus it had to be ascertained that any increase in  $\beta$ -galactosidase activity exhibited at the non-permissive temperature was not due to, for example, increased rates of transcription and/or translation, but due solely to changes in the post-transcriptional processing of pre-mRNA. Consequently an intron-less *MATa1-lacZ* gene fusion was engineered which encoded an active  $\beta$ -galactosidase fusion protein. This gene fusion could be used to monitor the effects of temperature on aspects of gene expression other than pre-mRNA processing.

Other yeast genes were considered, notably actin, but construction of an ORF in alternative introns would have been very time-consuming because most introns are long (300-500 nucleotides) and contain multiple stop codons in all reading frames.

### 3.3 PLASMID CONSTRUCTION

A DNA fragment encoding exon 1, intron 1 and part of exon 2 of the *MATa1* gene was sub-cloned into M13mp9, and SDM was performed on the cloned fragment to produce

various changes. All M13mp9 clones were plaque-purified three times (and sequenced each time), prior to isolation of the altered *MATa1* sequence as a *Bam*HI/*Eco*RI fragment. In a tri-molecular reaction the *MATa1* and *lacZ* genes were ligated into a yeast shuttle vector, pBM272 $\Delta$ RI, permitting generation of a fusion transcript initiating in, and under the control of, the *GALI* promoter. All plasmid and RF phage constructs were verified by restriction enzyme analysis, confirming the presence of individual fragments in single copy (results not presented; see Fig. 3.4).

### 3.3.1 Parent Plasmids

#### 3.3.1.1 pBEMAT/S

Plasmid pBEMAT/S (a gift from H. Domdey ; see Fig. 3.1) is the plasmid pSP64 carrying a modified *MATa1* gene in which intron 1 has been mutated by a three base-pair substitution to produce a *Sma*I site just upstream of the TACTAAC box. This also fortuitously changes the only stop codon in the frame that allows translational read-through of the intron. When compared to the normal sequence these mutations do not appear to affect the splicing efficiency of the intron either *in vitro* or *in vivo* (Kohrer and Domdey, 1988). Thus for our purposes this gene is wild-type and will be referred to as such in this thesis.

#### 3.3.1.2 pDEV19

This plasmid was a gift from M. Winther, Wellcome Biotechnology Ltd. London, and contains the coding sequence of the *lacZ* gene of *E.coli*. At the translation initiation site of this *lacZ* sequence is an *Eco*RI restriction site which is cleaved after the G nucleotide of the initiating ATG. The entire *lacZ* coding sequence (minus the initiating ATG) can be isolated as a 3.128kb *Eco*RI/*Sal*I fragment. At least the first 27 amino acids of  $\beta$ -galactosidase are not

required for activity (Casadaban *et al.*, 1980) therefore the absence of the initiating ATG does not affect activity as long as translation is initiated at an upstream, in-frame ATG (i.e. in a gene fusion construct) or at an ATG within the first 27 amino acids of *lacZ* (of which there are normally none; see above paper). Thus in any fusion construct utilising this *EcoRI/SalI lacZ* fragment, translation should initiate at the first ATG 5' to *lacZ* (if such an ATG is supplied as part of the amino terminal gene fusion). Within the space of eight codons downstream of the *lacZ EcoRI* site described above there are *SmaI* and *BamHI* sites present (in that order). Thus *SmaI/SalI* and *BamHI/SalI lacZ* restriction fragments will also encode active  $\beta$ -galactosidase proteins, subject to the same constraints as outlined for the *EcoRI/SalI* fragment described above. However, the *lacZ* gene on pDEV19 has an 8bp *ClaI* linker inserted at the *SmaI* site which unfortunately introduces an in-frame ATG between the *EcoRI* and *BamHI* sites. Thus *EcoRI/SalI lacZ* fragments from pDEV19 contain an in-frame ATG at the 5' of the *lacZ* coding sequence, whereas *BamHI/SalI* fragments do not. The presence of this in-frame ATG raises the possibility of internal initiation of translation, producing  $\beta$ -galactosidase activity from both a pre-mRNA and mRNA transcript (i.e. activity which is independent of whether sequences upstream, in this case the *MATa1* intron, are spliced or unspliced). Such activity would not satisfy criterion (a); see Section 3.2 above. pDEV19 was used as a source of the *lacZ* gene.

### 3.3.1.3 pSP*lacZ*

The *BamHI/SalI lacZ* fragment of pDEV19 was cloned into pSPT19, thus eliminating the internal ATG present between the *EcoRI* and *BamHI* sites of the *lacZ* coding sequence (see above) and facilitating the removal of *lacZ* as an *EcoRI/SalI* fragment. In this construct the *EcoRI/BamHI* region of the *lacZ* fragment isolated is entirely pSPT19 poly-linker sequence, but still produces an in-frame fusion without eliminating  $\beta$ -galactosidase activity. This plasmid was used as a source of the *lacZ* gene, and to distinguish this from that of pDEV19 it is termed

the *placZ* fragment. The difference between *lacZ* (pDEV19) and *placZ* was verified by restriction enzyme analysis (results not presented). The final yeast expression plasmids containing the *placZ* fragment are pJBM-5, pJBM-4, pJBM-4/s and pJBM-1/s. The plasmids pJBM-1 and 2 contain the *lacZ* fragment from pDEV19.

#### 3.3.1.4 pBM272 $\Delta$ RI

This plasmid is a derivative of pBM125 (a gift from R. Davies). It is a centromeric yeast shuttle vector containing the *GAL1/10* promoter region. Cloning into the *Bam*HI site of this vector puts the coding sequence of interest under the control of the *GAL1* promoter with translation initiating at the first ATG of the insert (in our case this was present on the *MATa1* component of the fusion with *lacZ*). The *Eco*RI site was removed to allow the creation of the final constructs (see Fig. 3.1).

#### 3.3.2 Steps in plasmid construction

All DNA fragments used in cloning experiments were individually gel purified (see Chapter 2.2.5.7), including the plasmid vector pBM272 $\Delta$ RI when cut with *Bam*HI/*Sal*I.

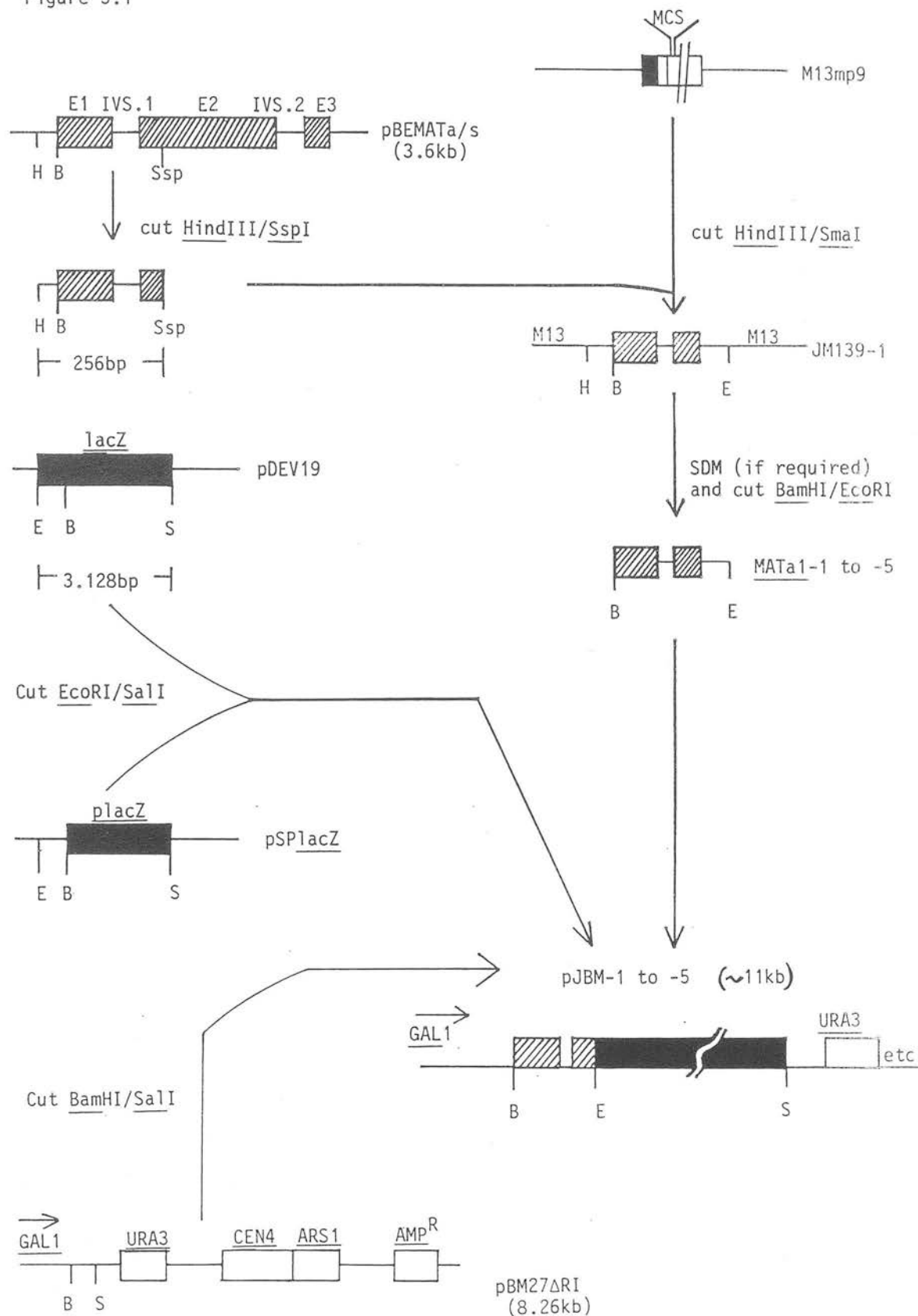
The source of the *MATa1* gene for all constructs was pBEMAT/S. The 256bp *Hind*III/*Ssp*I fragment containing E1, IVS-1 and 54 nucleotides of E2 was sub-cloned into M13mp9 cut with *Hind*III and *Sma*I (see Fig. 3.1). This fragment should have disrupted the reading frame of the *lacZ* gene present in M13mp9; however the colourmetric test used to determine if the insertion was successful produced only blue plaques, indicating failure to sub-clone. The presence of this fragment in M13mp9 was established by plaque hybridisation using the 256bp fragment as a probe (results not presented). The reason for the failure of this fragment to eliminate  $\beta$ -galactosidase activity in M13mp9 is unknown. Internal initiation of translation using the *MATa1* ATG (present in E1), or its use on a transcript originating within

### FIGURE 3.1: CONSTRUCTION OF PLASMIDS pJBM-1 TO pJBM-5

A flow diagram is depicted showing the plasmids and gene fragments used in the generation of the final yeast centromeric expression vectors (see Section 3.3).

*Bam*HI = B,    *Eco*RI = E,    *Hind*III = H,    *Sal*I = S,    *Ssp*I = Ssp

Figure 3.1



the *HindIII/SspI* cloned fragment are distinct possibilities, because the *MATa1* ATG is in frame with the M13mp9 *lacZ* gene. When cloned in the opposite orientation into M13mp18, this fragment did eliminate  $\beta$ -galactosidase activity. Sequencing in both orientations showed *MATa1* to be correct (see Fig. 3.4, only one orientation presented) i.e. it was wild-type. A series of mutations was then introduced into this gene (see section 3.3.3). Each mutated sequence was then isolated as a 232bp *BamHI/EcoRI* fragment from its respective M13 RF clone and ligated along with the *EcoRI/SalI lacZ* or *placZ* coding sequence into pBM272 $\Delta$ RI cut with *BamHI* and *SalI*. Thus a series of constructs was created that contain wild-type and mutant versions of a region of the *MATa1* gene fused to the *lacZ* gene and expressed under the control of the *GAL1* promoter. The mutations varied the  $\beta$ -galactosidase activity profile of the final plasmids as discussed in the next section.

### 3.3.3 Site-Directed Mutagenesis

#### 3.3.3.1 Mutant Phage Detection

A simplified screen for detection of plaques containing mutant phage was utilised. Phage containing specifically induced mutations are normally detected by hybridisation of the radioactively labelled mutating oligodeoxynucleotide to plaques immobilised on nylon membranes. By gradually increasing the wash temperature mutant plaques may be differentiated from those containing the parental phage (see below). This process involves autoradiography of the membranes after each washing stage and depending on the strength of the radioactive signal can extend over a period of days (due to prolonged exposure intervals).

Single-stranded M13 DNA, and M13 DNA containing the sequence to be mutated, were spotted onto a control membrane (CM) at various dilutions. The mutating oligodeoxynucleotide was then hybridised to it in the same manner as the plaques to be screened (mutant screen) i.e. at room temperature for 3-4 hours (see Chapter 2.2.5.16). Both types of membrane (the CM and the mutant screen membrane, MSM) were then washed in 6

X SSC at room temperature (3 x 100ml, 3 minutes each). The mutant screen membranes were then wrapped in cling film and stored at -20°C. The control membrane was washed at increasing temperatures (2 x 100ml, 3 minutes each: usually only two points) until about 5°C below its calculated  $T_m$ , after which 2-3°C increments were used. After each wash it was monitored with a Geiger counter for loss of radioactivity. In all cases the mutating oligodeoxynucleotide came off the unmutated M13 insert close to its calculated  $T_m$ . The temperature at which the oligodeoxynucleotide came off was determined and the mutant screen membranes were then washed at 2-4°C above this. Only those plaques which harboured phage containing the mutated sequence that was complementary to the mutating oligodeoxynucleotide remained radioactively labelled. The membranes were then exposed to X-ray film overnight (see chapter 2.2.1.4 and Fig. 3.2). By avoiding autoradiographic exposure after each wash, the time spent determining the temperature at which the mutant screen membranes had to be washed was reduced considerably.

### 3.3.3.2 Changes to *MATa1* Region in M13mp9

Tables 3.1 and 3.2 contain details of the oligodeoxynucleotides and templates used in the derivation of the various mutant constructs. Figures 3.3A and 3.4 show the *MATa1-1* translational reading frame and the mutated *MATa1* sequences respectively. Figure 3.3B shows the expected  $\beta$ -galactosidase activity profile of the mutant constructs, indicates the positions of internal ATG codons that could theoretically allow internal initiation of translation (see below) and shows the position of the engineered stop codons in the introns. It was assumed that the changes in the amino acid sequence caused by either splicing and/or specific mutations would have no effect on the  $\beta$ -galactosidase activity of the fusion. Each mutation introduced into the cloned *MATa1* fragment (see below) is described individually. The starting point is the wild-type *MATa1* region cloned into M13mp9. The *Bam*HI/*Eco*RI fragment isolated from this source is referred to as *MATa1-1* and the final construct is pJBM-1. This construct encodes  $\beta$ -galactosidase fusions from both pre-mRNA and mRNA transcripts.



- Mutation 1: a single base-pair substitution (A>T) was introduced to create the stop codon TAG in the intron reading frame. The pre-mRNA from the final construct should not produce an active fusion due to the block in translation. As a result of splicing this stop codon was not present on the mRNA and therefore it could encode  $\beta$ -galactosidase activity. The result was creation of the *Bam*HI/*Eco*RI *MATa1-1/s* fragment and the final construct, pJBM-1/s, was used in studies of pre-mRNA commitment to splicing (see Appendix).
- Mutation 2 : the *MATa1-1* region had to be altered because the intron contained a multiple of 3 nucleotides such that E1 and E2 were in the same translational phase in both spliced and unspliced RNA. It was required that splicing produce a mRNA which could not encode an active  $\beta$ -galactosidase fusion. To fulfill this specification the intron was lengthened by 5 bases into E2 which was reciprocally reduced. This meant that the intron was no longer a multiple of three and its removal caused a change in the reading frame with the result that an active fusion was encoded by unspliced but not spliced transcript. This increase in the length of the intron was achieved by mutating the 3' splice site and producing a new one 5 bases into E2. In the resulting *MATa1-2* fragment the length and reading frame of the pre-mRNA remained unaltered.
- Mutation 3 : an ATG>ATC mutation was introduced just downstream of the new 3' splice site. This removed an in-frame initiation codon which could possibly lead to internal initiation of translation on the mRNA (and pre-mRNA) encoded by the *MATa1-2* mutant fragment causing background problems in the screening procedure. The *MATa1-3* mutant fragment was thus produced.

- Mutation 4: an out of frame ATG, prior to the intron, which in *MATa1-2* and *MATa1-3* spliced back into frame with no intervening stop codons, was mutated to ACG. This resulted in the *MATa1-4* screening fragment which was used in the final screening construct pJBM-4.
- Mutation 5: using SSM139-4 (see Table 3.2) as the template, a stop codon was introduced into the intron (see *MATa1-1/s* above). The resulting sequence, *MATa1-4/s*, should not produce  $\beta$ -galactosidase from any encoded RNA in the final construct.
- Mutation 6: the intron of the *MATa1-4* fragment was completely removed to give *MATa1-4 $\Delta$ int*. This was equivalent to an accurately spliced transcript from construct pJBM-4 which should not encode  $\beta$ -galactosidase activity.
- Mutation 7 : using SSM139-4 $\Delta$ int as the template, the first nucleotide of the new E2 was deleted to generate an intronless in-frame *Bam*HI/*Eco*RI *MATa1-5* fragment. An in-frame  $\beta$ -galactosidase fusion would be produced from the final construct pJBM-5. This is the control for changes in the rates of RNA processing events other than those associated with pre-mRNA.

The  $\beta$ -galactosidase activity profile expected of the final constructs, pJBM-1 to pJBM-5, is given in Fig. 3.3B and in the summary below.

### 3.4 SUMMARY

To develop a new screening procedure for the identification of conditional lethal *prp* mutants a series of novel gene fusions had to be created. Part of the *MATa1* gene of yeast was fused to the *lacZ* gene of *E.coli* and placed into a yeast expression vector under the control of the *GAL1* promoter. By introducing specific mutations into the *MATa1* region, fusion proteins were produced which were expected to encode  $\beta$ -galactosidase activity as indicated below. Three of these final constructs (marked \* below) were used to ascertain whether the proposed screen was viable (see Chapter 4), whilst all 6 were used in the study of pre-mRNA commitment to splicing (see Appendix).

#### $\beta$ -galactosidase activity

<u>plasmid</u>	<u>pre-mRNA</u>	<u>mRNA</u>
pJBM-5 (*)		+
pJBM-1	+	+
pJBM-1/s	-	+
pJBM-2	+	-
pJBM-4 (*)	+	-
pJBM-4/s (*)	-	-

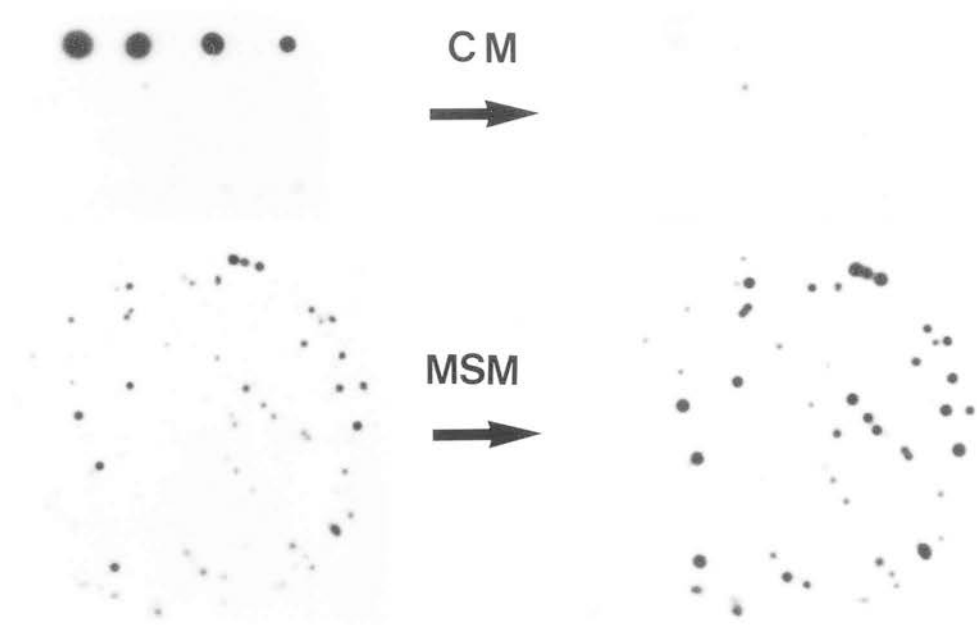
/s plasmids have a stop codon in the intron reading frame.

### FIGURE 3.2: WASHING PROCEDURE FOR THE DETECTION OF MUTANT PHAGE

The hybridisation and washing procedures utilised in this simplified screening procedure, designed to detect mutant phage-containing plaques, are described in Section 3.3.3.1. Single-stranded M13 DNA (insert-containing and insert-minus) was spotted onto the control membrane (CM) at varying concentrations (0.25 $\mu$ g, 0.5 $\mu$ g, 1 $\mu$ g, 2 $\mu$ g). The DNA was fixed to the membrane and a radioactively labelled oligodeoxynucleotide was hybridised to the DNA. The CM was then washed in 6 X SSC at increasing temperatures and monitored after each wash by a Geiger counter. After washing most of the radioactivity off the CM, the temperature was raised 2-3°C and the mutant screen membrane (MSM) was washed at this increased temperature. This figure shows the results normally obtained using this procedure.

The M13 insert-containing DNA is identified on the CM by its ability to hybridise with the oligodeoxynucleotide (see the heavy radioactive signal on the CM and note that this decreases along the membrane due to the reduced DNA concentrations of the individual spots). The insert-minus M13 DNA (spotted just below the insert-containing M13 DNA) fails to hybridise with the oligodeoxynucleotide and therefore produces only a non-specific background signal.

FIG. 3.2



### FIGURE 3.3A: THE *MATa1* SEQUENCE AND READING FRAME

The *MATa1* (*MATa1-1*; see Fig. 3.1) sequence from pBEMATa/S is given from the *Bam*HI (B) to *Ssp*I (Ssp) sites. Nucleotides which have been mutated in the generation of the other *MATa1* sequences (see Table 3.2) are indicated in bold-type. Exon and intron sequences are in capital and lower case letters respectively. The translational reading frame of the *MATa1-lacZ* fusion is indicated by the vertical lines above and between the nucleotides. The internal ATG codons (see Section 3.3.3.2) are underlined, though the ATG codon present at the 5' end of the *lacZ* *Eco*RI/*Sal*I fragment (see Sections 3.3.1.2 and 3.3.1.3) is not shown.

### FIGURE 3.3B: THE EXPECTED $\beta$ -GALACTOSIDASE ENCODING ABILITIES OF THE *MATa1-lacZ* TRANSCRIPTS OF pJBM-1 TO pJBM-5

The transcripts encoded by the various *MATa1-lacZ* fusions are shown. The positions of the internal AUG codons and stop codons (within the intron) are indicated by an A and S respectively. The expected  $\beta$ -galactosidase coding ability of the transcripts are represented by the (+) and (-) signs, where (+) indicates that they have the potential to encode an active  $\beta$ -galactosidase fusion protein.

### FIGURE 3.3C: A FLOW DIAGRAM OF THE STEPS INVOLVED AND THE OLIGONUCLEOTIDES USED IN THE CREATION OF THE VARIOUS *MATa1* CONSTRUCTS OBTAINED BY SDM

Figure 3.3C shows the derivation of the various *MATa1* sequences created by successive mutageneses of various single-stranded (SS) M13 templates which contain the specified *MATa1* sequences. The starting template is SSJM139-1, which gives rise to the *MATa1-1* fragment (see Table 3.2). All mutant templates were derived from the SS M13 template immediately prior to them in this flow diagram (as indicated by the arrows). Below each SS template e.g. SSJM139-1, is the name of the final yeast centromeric expression vector in which the specific *MATa1* sequences reside e.g. pJBM-1, and also whether these plasmids contain the *lacZ* or *placZ* fragments (see Section 3.3.1.4, p94).

The changes to the 3' splice site made by Mutation 2 are also indicated. Intron and exon sequences are in lower and upper case letters respectively.

Figure 3.3A

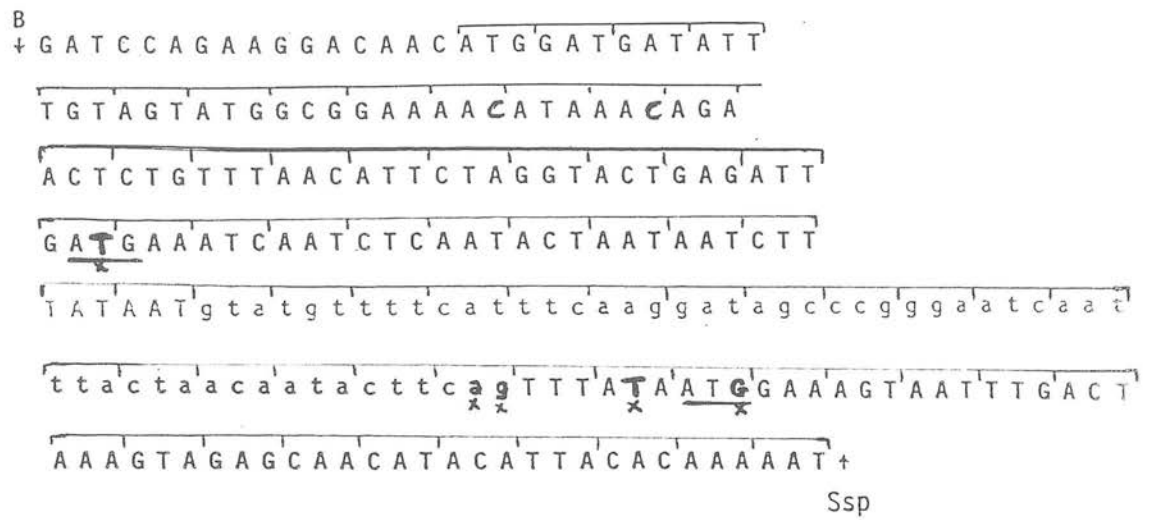


Figure 3.3B

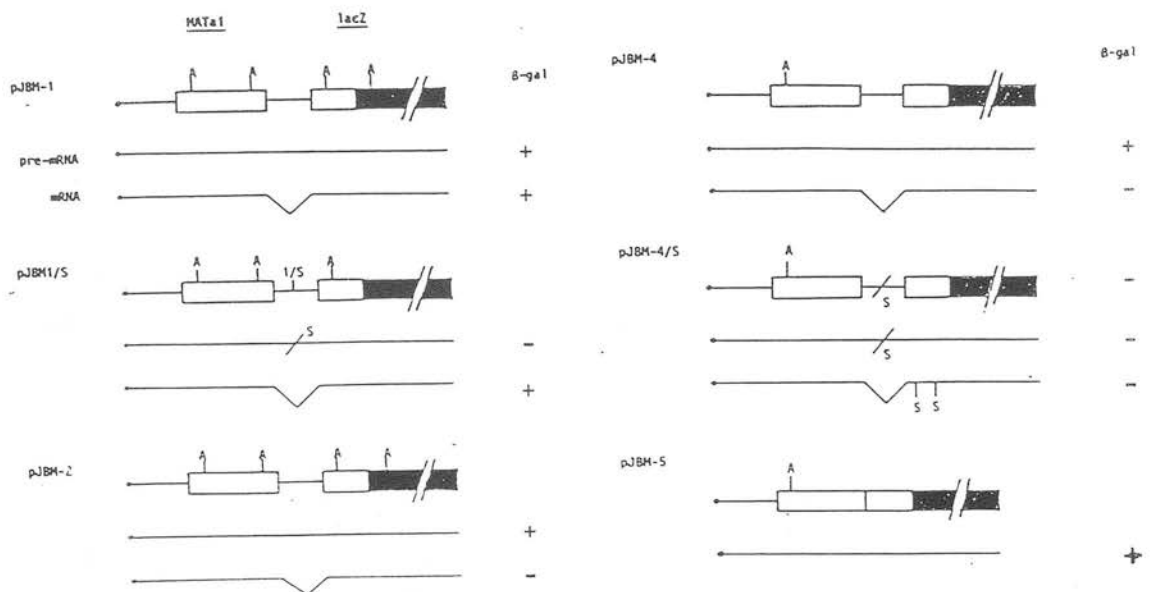
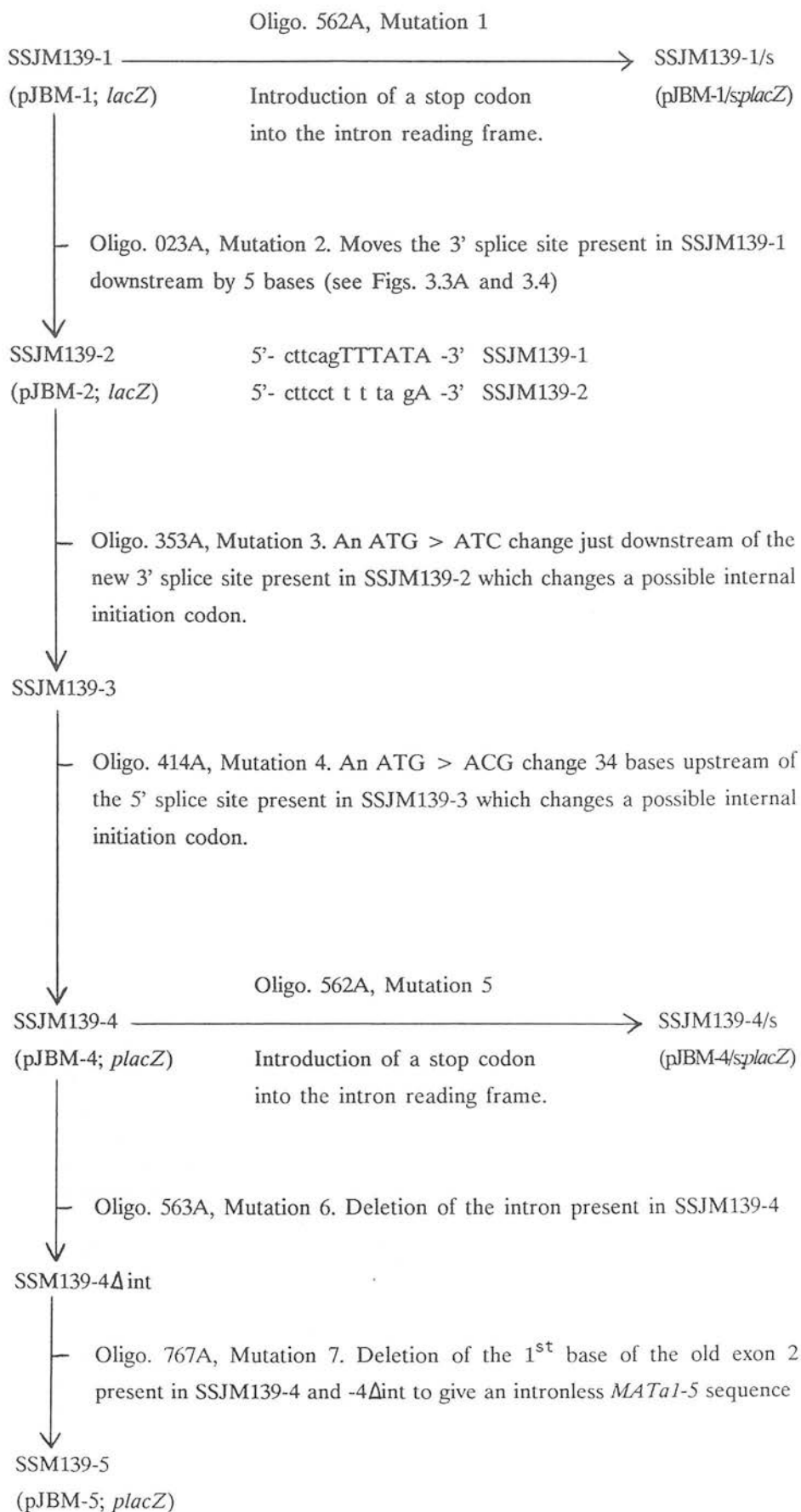


FIGURE 3.3C





**TABLE 3.1** **OLIGODEOXYNUCLEOTIDES USED IN SITE-DIRECTED MUTAGENESIS**

Oligonucleotide Name	Oligonucleotide Sequence (5' → 3')	Mutation (see Section 3.3.3.2)
562A (JM05)	GCTATCCTAGAAATGAA	1 and 5
023A	TTTCATTCTAAAAGGAAGTATTG	2
353A (JM03)	TTACTTTCGATTCTAAA	3
414A (JM04)	GATTTCGTCAATCTCAG	4
563A (JM06)	ACTTTCGATTATTATAAAGA	6
767A (JM07)	CTTTCGATATTATAAAG	7

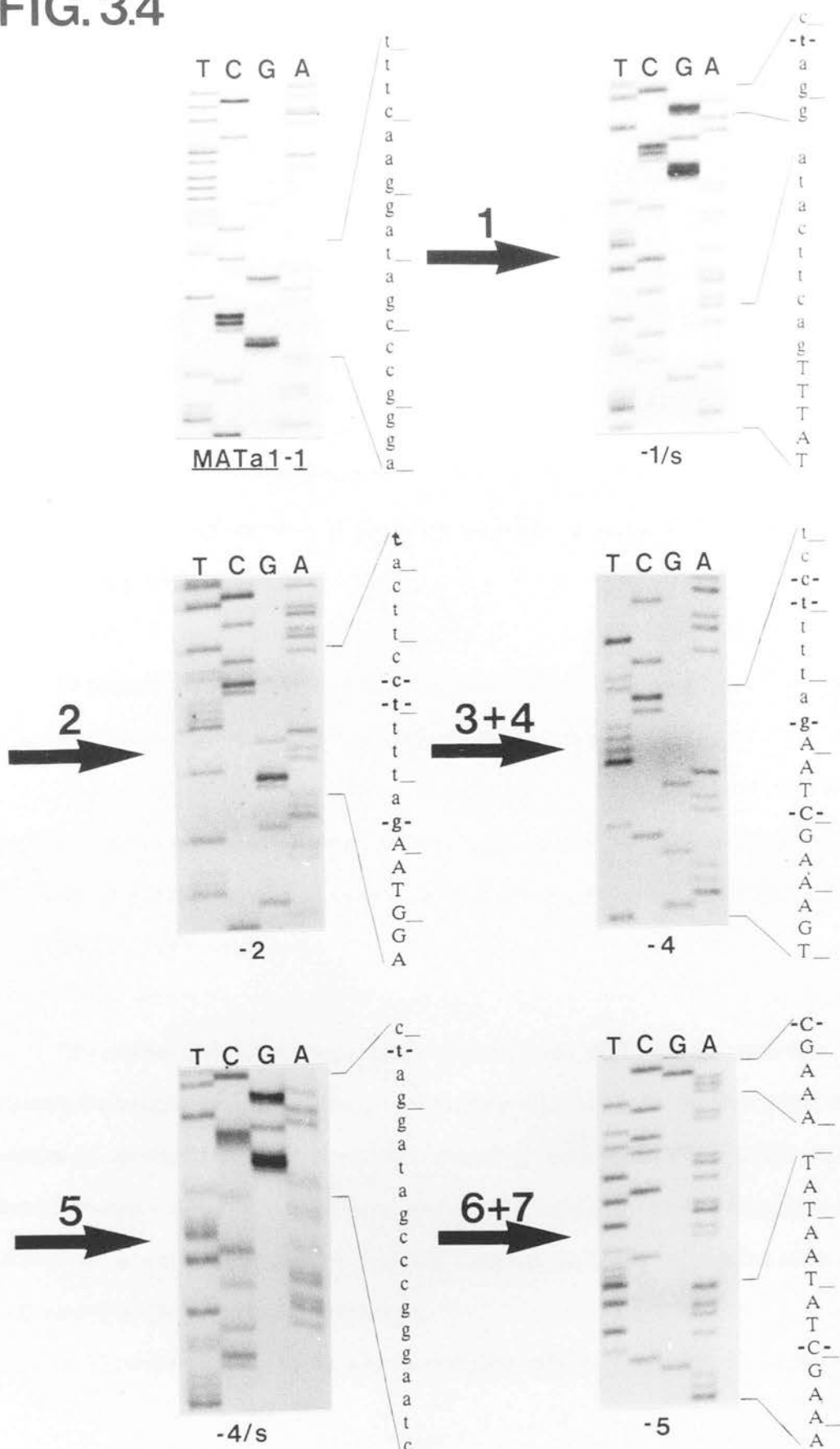
**TABLE 3.2** **SDM: TEMPLATES, FRAGMENTS AND FINAL CONSTRUCTS**

Template	Mutation	Single-Stranded M13 Mutant Template	Isolated M13 BamHI/EcoRI Fragment	Final Plasmid Construct
SSM139-1	-	-	<u>MATa1-1</u>	pJBM-1
SSM139-1	1	SSM139-1/s	<u>MATa1-1/s</u>	pJBM-1/s
SSM139-1	2	SSM139-2	<u>MATa1-2</u>	pJBM-2
SSM139-2	3	SSM139-3	-	-
SSM139-3	4	SSM139-4	<u>MATa1-4</u>	pJBM-4
SSM139-4	5	SSM139-4/s	<u>MATa1-4/s</u>	pJBM-4/s
SSM139-4	6	SSM139-4Δint	-	-
SSM139-4Δint	7	SSM139-5	<u>MATa1-5</u>	pJBM-5

#### FIGURE 3.4: GENERATION OF THE VARIOUS MUTANT *MATa1* SEQUENCES

The mutated *MATa1* sequences produced by site-directed mutagenesis were generated in a stepwise manner (see Section 3.3.3.2). The steps involved and the nucleotide changes made are shown in this figure; the numbered arrows correspond to the mutations described in Section 3.3.3.2. The mutated nucleotides (i.e. those changed from the original *MATa1-1* sequence) are indicated by the horizontal bar lines positioned on either side of them. Exon and intron sequences are given in capital and lower case letters respectively. The reading frames of the various *MATa1-lacZ* fusions are shown by the faint horizontal lines separating the nucleotides.

FIG. 3.4



## CHAPTER 4

### FEASIBILITY AND DEVELOPMENT OF THE SCREENING PROCEDURE

#### 4.1 THE FEASIBILITY OF A *prp* MUTANT SCREEN

Using the plasmids pJBM-5, -4 and -4/s, the feasibility of detecting *prp* mutants by means of a postulated increase in their  $\beta$ -galactosidase activity at the non-permissive temperature was examined. There were a number of questions concerning this proposition that needed to be addressed.

- (a) Was the fusion-transcript generated from the screening construct pJBM-4 spliced at the new 3' splice site under all conditions (see Chapter 3.3.3.2 and compare pJBM-1 with pJBM-4)?
- (b) Did pre-mRNA, generated from pJBM-4, accumulate in a *prp* strain when splicing was inhibited i.e. after incubation at the non-permissive temperature?
- (c) Were increases in  $\beta$ -galactosidase activity caused by pre-mRNA accumulation rather than by increases in the rates of other RNA processing events e.g. transcription, RNA transport and/or translation?

Experiments to resolve these questions were performed at the permissive and non-permissive temperatures because the screen involved the incubation of strains under these conditions. In addition it had to be ascertained whether the wild-type strain utilised in making *r<sup>s</sup>* mutants complied with a certain condition relevant to the feasibility of the screen; it should not accumulate pre-mRNA (hence  $\beta$ -galactosidase) at the non-permissive temperature above that already present at the permissive temperature.

To address the questions above, a series of experiments was performed scrutinising both

the fusion transcript and  $\beta$ -galactosidase activity profiles emanating from pJBM-4 (the screening construct).  $\beta$ -galactosidase activities only were assayed for the control constructs pJBM-5 and pJBM-4/s.

#### 4.1.1 Primer Extension Analysis of the *MATa1-lacZ* Fusion-Transcript from pJBM-4 in Wild-Type and *prp* Strains

The RNA profile of the fusion-transcript generated from pJBM-4 was investigated by primer extension analysis using an oligodeoxynucleotide complementary to the 5' end of the *lacZ* coding sequence (see Chapter 2.1.6 and Table 2.4). The wild-type strain, DBY745, was compared to a *prp2* mutant, DJY36, at the permissive (23°C) and non-permissive (36°C) temperatures (see Fig. 4.1). Splicing of the wild-type *MATa1* transcript is normally inefficient (Miller, 1984; Ner and Smith, 1989), producing a ratio of roughly 20% pre-mRNA : 80% mRNA. Densitometric scanning of primer-extended products from DBY745 (pJBM-4) indicated that the pJBM-4 fusion-transcript was spliced at roughly 72%-77% efficiency at the permissive temperature (results not presented). Taking into consideration the probable differences in splicing efficiency of unrelated strains, this figure suggests that splicing of the pJBM-4 transcript is normal and that changing the position of the 3' splice site had little or no effect on splicing efficiency (see Fig. 4.1A)

The first AG downstream of the TACTAAC box is used as the 3' splice site in yeast (Fouser and Friesen, 1987) and this AG is normally preceded by a pyrimidine nucleotide. When this pyrimidine is altered to a purine the use of cryptic 3' splice sites (at very low efficiency) has been detected (Langford and Gallwitz, 1983). The next AG downstream of the new 3' splice site (on the *MATa1-4* part of the fusion-transcript generated from pJBM-4), which would allow splicing back into frame and therefore enable the mRNA to encode  $\beta$ -galactosidase activity, was preceded by a purine nucleotide. It seemed likely that the new 3' splice site (which was preceded by a pyrimidine nucleotide) would be selected in preference to all other potential sites. Deletion analysis has shown that for efficient splicing there is both a minimum and a

maximum distance requirement between the TACTAAC box and the 3' splice site (roughly 8-66nts) (Langford and Gallwitz, 1983; Cellini *et al.*, 1986a). It should be noted that *MATa1* IVS-1 has only 10 nucleotides separating the TACTAAC box from the 3' splice site, and this may be one reason for its inefficient splicing. The new 3' splice site of pJBM-4 was the first AG downstream of the TACTAAC box, and increased the distance from the branch point by only 5 bases; thus it was considered that this would have little (if any) effect on the choice of the 3' splice site. Comparison of the sizes of the major primer extension products indicated that the new 3' splice site was the only one utilised efficiently (see Fig. 3.4 and 4.1).

Thus in both a wild-type (DBY745) and a *prp2* (DJY36) strain incubated at 23°C the pJBM-4 fusion-transcript is spliced at the new 3' splice site with similar efficiency to that seen for the normal *MATa1* IVS-1 (see Fig. 4.1A and C). It is also apparent from Fig. 4.1B that in DBY745 incubated at 36°C, the fusion-transcript was spliced at the new 3' splice site and there was no accumulation of pre-mRNA above that exhibited on incubation at 23°C. Indeed splicing appears to be more efficient at 36°C in DBY745 i.e. the ratio of pre-mRNA : mRNA decreases. Densitometric scanning of the primer-extended products generated from the fusion-transcript encoded by pJBM-4 indicates that the fusion-transcript is spliced with roughly 85% efficiency when cultures of DBY745 (pJBM-4) are incubated at 36°C (results not presented).

When analysed in a *prp2* mutant at the non-permissive temperature (36°C, pJBM-4 fusion-transcript pre-mRNA accumulated (relative to its mRNA and internal PGK control), indicating that the intron from this fusion is not removed when splicing is inhibited (see Fig. 4.1D). Thus it can be concluded that the screening construct pJBM-4 behaves as proposed i.e. the pre-mRNA of the fusion-transcript is spliced at the new 3' splice site with wild-type efficiency when cultures are incubated at 23°C, and accumulates when splicing is inhibited i.e. when *prp* mutant cultures are incubated at 36°C.

An oligodeoxynucleotide (JMO-PGK) complementary to a region of the intron-less phosphoglycerate kinase (*PGK*) transcript (see Chapter 2.1.6 and Table 2.4) was used as an internal control to allow monitoring of relative RNA levels in each primer extension reaction.

**FIGURE 4.1: PRIMER EXTENSION ANALYSIS OF THE *MATa1-4* FUSION  
TRANSCRIPT IN WILD-TYPE AND *prp2* STRAINS**

All primer extension analyses (see Chapter 2.2.8.5) described in this figure, utilised the oligodeoxynucleotides 083C (which hybridises to the 5' end of the *lacZ* region of the *MATa1-lacZ* fusion transcript), and JMO-PGK (which hybridises to the intron-less *PGK* transcript: see Chapter 2.1.6 and Table 2.4). The sizes of the primer-extended species produced by the hybridisation of the oligodeoxynucleotides to specific RNAs are given below.

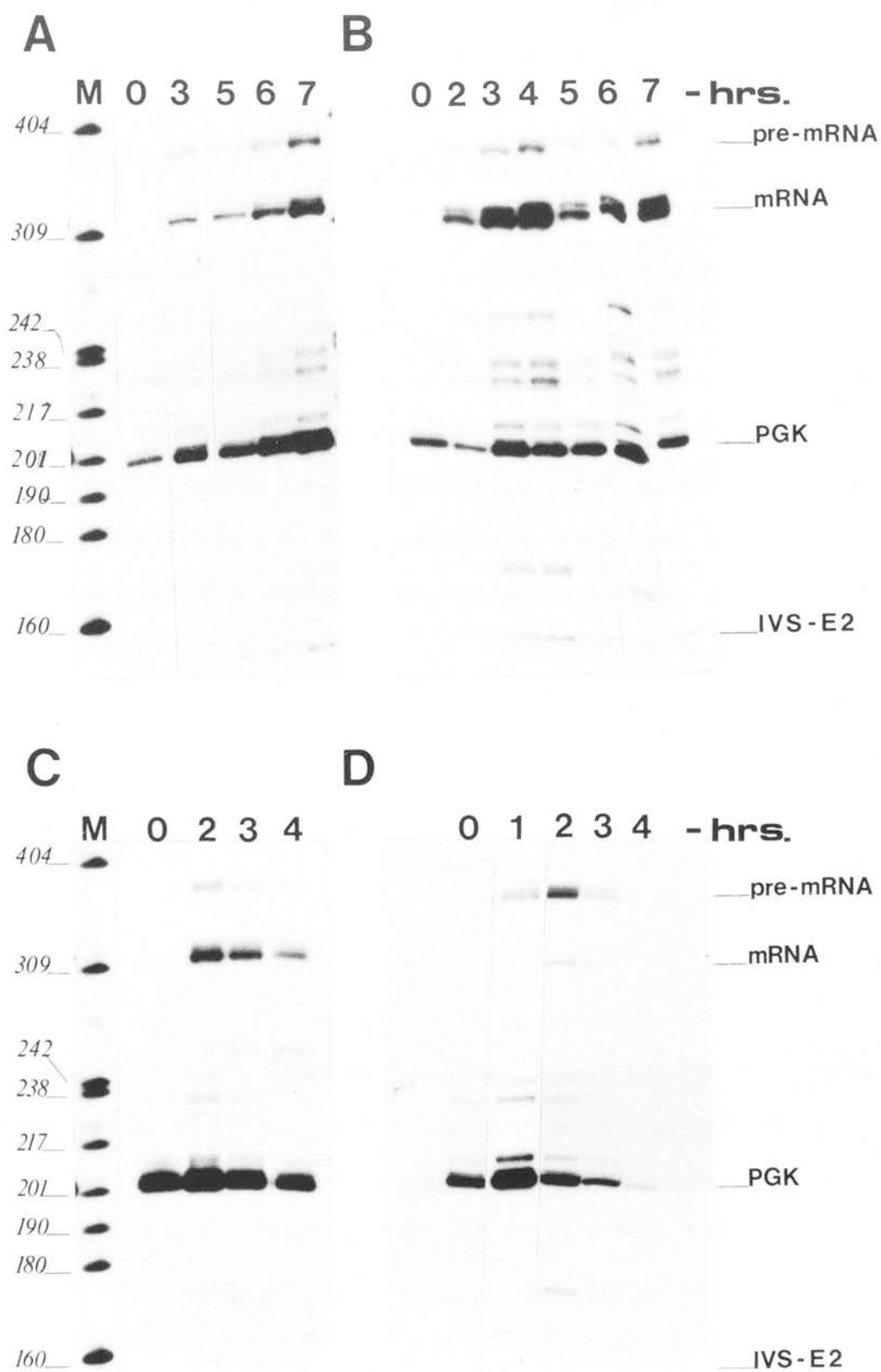
083C - pre-mRNA - 381nts  
mRNA - 322nts  
IVS-E2 - 155nts (083C does not hybridise to the IVS RNA)  
JMO-PGK - mRNA - 200nts

The primer extension product of JMO-PGK is denoted as PGK. All other species indicated arise from 083C and therefore represent the *MATa1-4-lacZ* fusion transcript molecules generated from pJBM-4. The identification of the IVS-E2 intermediate (of *MATa1-4-lacZ*) is made on the basis that the presumed primer extension product is the correct size and is absent in the analysis of a *prp2* strain incubated at 36°C only (see Fig. 4.1D). The size markers are derived from a *MspI* digest of pBR322, the fragments of which have been rendered blunt-ended with radio-labelled nucleotides.

RNA preparations were made from galactose induced cultures as described in Chapter 2.2.8 and 2.2.9.1. The numbers above the tracks indicate the durations of galactose induction (in hours) before RNA was extracted from the cultures.

- A. Primer extension analysis of the *MATa1-4-lacZ* fusion transcript in the wild-type strain, DBY745, incubated at 23°C
- B. Primer extension analysis of the *MATa1-4-lacZ* fusion transcript in the wild-type strain, DBY745, incubated at 36°C
- C. Primer extension analysis of the *MATa1-4-lacZ* fusion transcript in the *prp2* strain, DJY36, incubated at 23°C
- D. Primer extension analysis of the *MATa1-4-lacZ* fusion transcript in the *prp2* strain, DJY36, incubated at 36°C

**FIG. 4.1**





It should be noted that the PGK mRNA shows a transient 6-7 fold accumulation caused by the transient heat-shock effect (Piper *et al.*, 1986). This effect lasts for 2 hours, then the RNA level returns to the pre-heat- shock state and must be taken into account when referring to RNA levels at the 1 hour time points at 36°C (see track 1, Fig. 4.1D).

#### 4.1.2 Analysis of Plasmid Encoded $\beta$ -galactosidase Activity

All experiments utilised two independent transformants for each plasmid (colony purified 3 times), which were both assayed in duplicate. The results presented are an average of both transformants. In general the transformants exhibited similar activity profiles and the standard deviations for the results observed for pJBM-5 were usually within 15% (results not presented). However individual pJBM-4-transformants of the *prp8* mutant strain did exhibit significant quantitative differences in  $\beta$ -galactosidase activity, though qualitatively they were very similar; see below. For a detailed protocol intimating the method of galactose induction of the transformants, please refer to Chapter 2.2.9. In this section cultures of transformants will be described thus; the name of the transformed strain e.g. DBY745, followed by the name of the plasmid, in brackets, residing in the strain e.g. (pJBM-5).

##### 4.1.2.1 DBY745

###### (a) pJBM-5 (intron-less control)

Figure 4.2 compares  $\beta$ -galactosidase activities measured from cultures of DBY745 (pJBM-5) incubated at 23°C and 36°C. It is clear that no significant difference in  $\beta$ -galactosidase activity exists between cultures incubated at the two temperatures until the 4 hour time-point where there is only a 1.25-fold decrease at 23°C relative to 36°C. This suggests that factors other than those involved in pre-mRNA processing e.g. rates of transcription/translation/RNA stability/transport and protein stability, remain similar in the shift

from 23°C to 36°C. It also implies that DBY745 is an ideal strain in which to recover *r<sup>s</sup> prp* mutants, as any substantial increase in  $\beta$ -galactosidase activity exhibited by the mutants on incubation at 36°C would probably be due to pre-mRNA processing defects.

(b) pJBM-4 (screening construct)

Figure 4.3 clearly indicates that cultures of DBY745 (pJBM-4) produce a measureable amount of  $\beta$ -galactosidase activity. The activity produced is consistently greater from cultures incubated at 23°C (about 1.5-fold greater) and is in contrast to the activity profile exhibited by DBY745 (pJBM-5) incubated at the two temperatures (see Fig. 4.2). This observation probably indicates a greater efficiency of pre-mRNA splicing at 36°C which would produce relatively less pre-mRNA and therefore less  $\beta$ -galactosidase fusion-protein; this is supported by densitometric scanning of primer-extended products generated from the fusion-transcript of pJBM-4 (see above). The fact that  $\beta$ -galactosidase was detected in DBY745 (pJBM-4) cultures means that pre-mRNA "leaks" out of the nucleus into the cytoplasm, thus producing a background level of activity. It was found that pJBM-4 produced roughly 2.3% of the  $\beta$ -galactosidase activity generated by pJBM-5 (values of 3-5 hour time-points added and averaged), and this was in good agreement with published results from similar experiments (Legrain and Rosbash, 1989; see the Appendix).

(c) pJBM-4/s (stop codon in intron)

In time-course induction experiments only the 5 and 6 hour time-points at 23°C and 36°C were assayed (results not presented). It was not expected that pJBM-4/s would generate much activity; however at 23°C it produced roughly 44% of that generated by pJBM-4, whereas at 36°C it produced 22%. pJBM-4/s did not produce such high percentage values of pJBM-4 in any other strain tested at either temperature (see 4.1.2.2 to 4.1.2.5 and 4.2.3) and therefore this observation appeared to be strain-dependent phenomenon. There are two possible

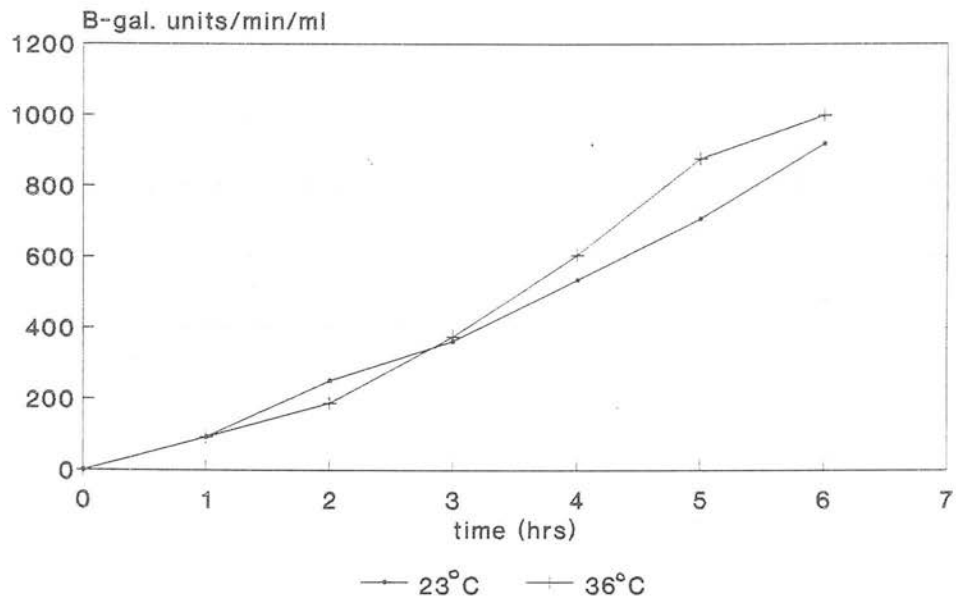
**FIGURE 4.2:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DBY745 (pJBM-5)  
INCUBATED AT 23°C AND 36°C: A WILD-TYPE STRAIN ANALYSIS**

Cultures were grown at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 prior to galactose induction and heat-shock (see Chapter 2.2.9). Culture-samples (0.5ml) were retrieved at hourly intervals after galactose induction, and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately after the addition of galactose, the so-called time-zero (see Chapter 2.2.9.1), and therefore 0 hours induction has an activity value of 0 (see Fig. 4.2).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

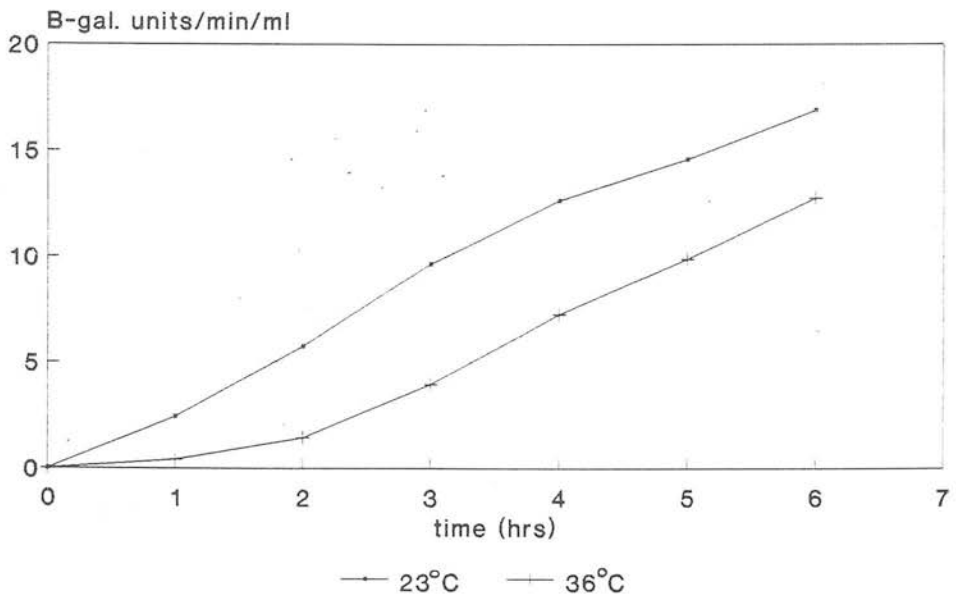
**FIGURE 4.3:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DBY745 (pJBM-4)  
INCUBATED AT 23°C AND 36°C; A WILD-TYPE STRAIN ANALYSIS**

All experimental details were as described for DBY745 (pJBM-5) in Fig. 4.2, except 0.5-2ml culture-samples were taken.

**FIG. 4.2**  
DBY745 (pJBM-5)



**FIG. 4.3**  
DBY745 (pJBM-4)



explanations that can account for the  $\beta$ -galactosidase activity profile exhibited by pJBM-4/s-containing cultures:

- (i) splicing at alternative splice sites on the fusion-transcript generated from pJBM-4/s restores the reading-frame of the *MATa1-lacZ* fusion. Results of primer extension analysis (see Section 4.1.1) argue against this possibility because the use of these sites (with such high efficiency) was not detected on the fusion-transcript generated from pJBM-4.
- (ii) translational read-through of the stop codon present in the intron of the pJBM-4/s fusion-transcript could generate  $\beta$ -galactosidase activity. There is some evidence to suggest that such translational read-through is possible in yeast (Brown, 1989).

#### 4.1.2.2 DJY36 (*prp2*)

##### (a) pJBM-5

Figure 4.4 compares the activity measured from cultures of DJY36 (pJBM-5) incubated at 23°C and 36°C. In contrast to DBY745 (results above), pJBM-5-generated activity in DJY36 fell about 10-fold when the cultures were incubated at 36°C compared to 23°C. This is probably a reflection of the fact that essential intron-containing genes (e.g. actin) are not being spliced at the non-permissive temperature resulting in cell stress and, eventually, cell death. This would lead to the decrease of  $\beta$ -galactosidase activity, due to the cessation of those events required to produce a functional  $\beta$ -galactosidase fusion-protein e.g. transcription/RNA processing/translation. However, it could be that DJY36 is *r<sup>s</sup>* for galactose induction and therefore cannot produce large amounts of fusion-transcript at the high temperature.

##### (b) pJBM-4

In Fig. 4.5 it can be seen that there was 3-fold greater activity exhibited by cultures incubated at 36°C compared to those incubated at 23°C. The increase in activity may be a

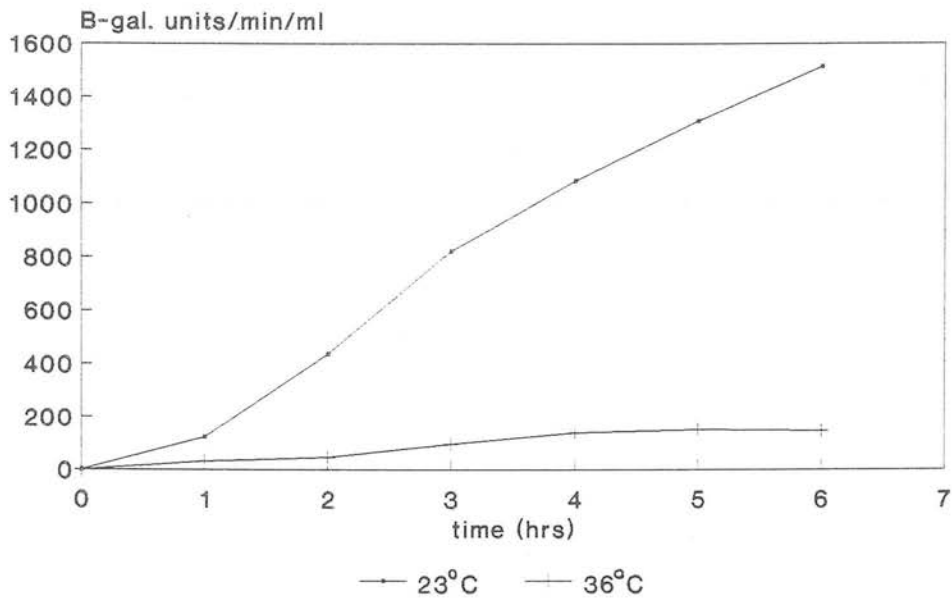
**FIGURE 4.4:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DJY36 (pJBM-5)  
INCUBATED AT 23°C AND 36°C: A *prp2* STRAIN ANALYSIS**

Cultures were grown at 23°C to OD 600<sub>nm</sub> of 0.1-0.4 prior to galactose induction and heat-shock (see Chapter 2.2.9). Culture-samples (0.5ml) were taken at hourly intervals after galactose induction, and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately after the addition of galactose, the so-called time-zero (see Chapter 2.2.9.1), and therefore 0 hours induction has an activity value of 0 (see Fig. 4.4).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

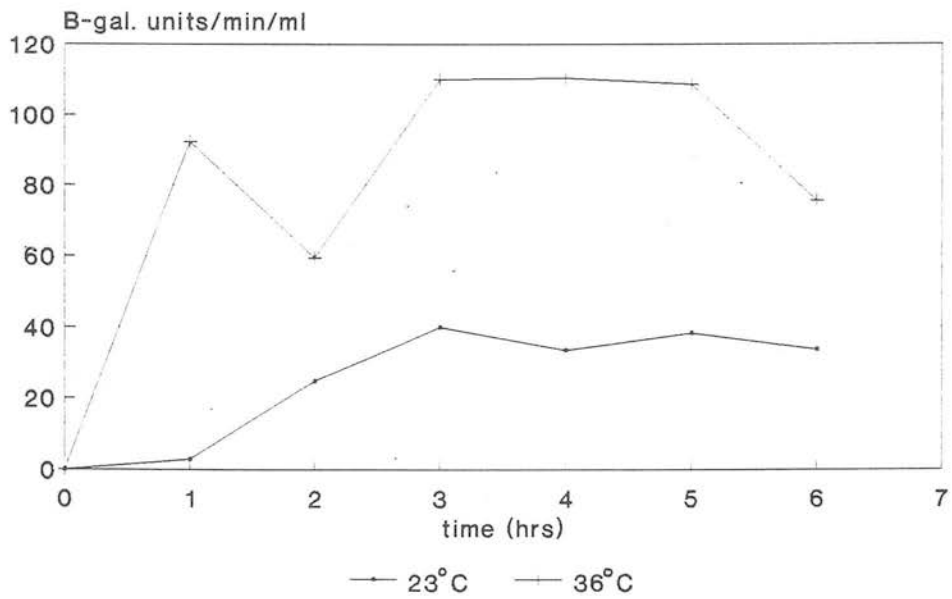
**FIGURE 4.5:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DJY36 (pJBM-4)  
INCUBATED AT 23°C AND 36°C; A *prp2* STRAIN ANALYSIS**

All experimental details were as described for DJY36 (pJBM-5) in Fig. 4.4, except 0.5-2ml culture-samples were taken.

**FIG. 4.4**  
DJY36 (pJBM-5)



**FIG. 4.5**  
DJY36 (pJBM-4)



result of the accumulation of pre-mRNA, which is a direct consequence of the *prp2* defect. The initial peak in activity observed at 36°C (1 hour time-point) was reproducible, and must be related to splicing as pJBM-5 did not demonstrate this effect (see Fig. 4.4). The reason for this sharp rise in  $\beta$ -galactosidase activity is not known and is specific to DJY36, though other (unassayed) *prp2* strains (MGL-1 and JBY27) do show a transient increase in viable cell number during the same period (results not presented). It is possible that the phenomena are associated. In DJY36 cultures incubated at 23°C (when splicing is functional), pJBM-4 generated roughly 3.5% of the activity originating from pJBM-5. This observation was close to that previously measured by Legrain and Rosbash though they only compared values obtained from a wild-type strain.

(c) pJBM-4/s

$\beta$ -galactosidase activities measured from cultures of DJY36 (pJBM-4/s) incubated at 23°C were 1.85% of those found in cultures of DJY36 (pJBM-4) (averaging the 4 and 5 hour time-points; results not presented). This suggests that read-through and/or alternative splicing are significantly reduced in this strain, and represents a >20-fold reduction compared to DBY745 (pJBM-4/s) incubated at the same temperature (see Section 4.1.2.1). More significantly, no activity was detected at 36°C. Both these observations suggest that the increase in  $\beta$ -galactosidase activity exhibited by cultures of DJY36 (pJBM-4) incubated at 36°C is due almost entirely to pre-mRNA accumulation.

4.1.2.3 SPJ8.31 (*prp8*)

(a) pJBM-5

Figure 4.6 compares activities exhibited by cultures of SPJ8.31 (pJBM-5) when incubated at 23°C and 36°C. Although the absolute  $\beta$ -galactosidase levels were very much



reduced, the overall activity profile generated by SPJ8.31 (pJBM-5) was similar to that seen for DJY36 (*prp2*)(pJBM-5). There was an 8-fold reduction in activity measured from cultures incubated at 36°C compared to 23°C and this is probably a consequence of cell death (see Section 4.1.2.2).

(b) pJBM-4

Figure 4.7 compares activities exhibited by cultures of SPJ8.31 (pJBM-4) when incubated at 23°C and 36°C. There was a 10-fold increase in activity measured at 36°C relative to 23°C, however the two transformants of this strain exhibited activities that were quantitatively, significantly different, though they were qualitatively similar (results not presented). In similar experiments Legrain and Rosbash observed a 3-4 fold increase in the same strain. The increase in activity was probably due to the accumulation of pre-mRNA in this *prp8* strain incubated at the non-permissive temperature. Cultures of SPJ8.31 (pJBM-4) incubated at 23°C produced roughly 1% of the  $\beta$ -galactosidase activity measured in cultures of SPJ8.31 (pJBM-5) under the same conditions. This is slightly less than that measured from DJY36 and DBY745 (3.5% and 2.3% respectively).

(c) pJBM-4/s

No activity arising from this plasmid was detected under any conditions. This confirmed that the activity measured from pJBM-4 was due solely to the pre-mRNA, and that the increase in activity exhibited by cultures of SPJ8.31 (pJBM-4) incubated at 36°C was a result of pre-mRNA accumulation caused by the splicing defect in this *prp8* mutant.

Similar analyses to those described above using other *prp* mutant strains proved to be impossible as the *prp3*, *prp4*, *prp5*, *prp6* and *prp11* strains in this laboratory were either *t<sup>s</sup>* for galactose induction (and therefore could not generate the pJBM-4 fusion-transcript at 36°C) or were not inducible at all.

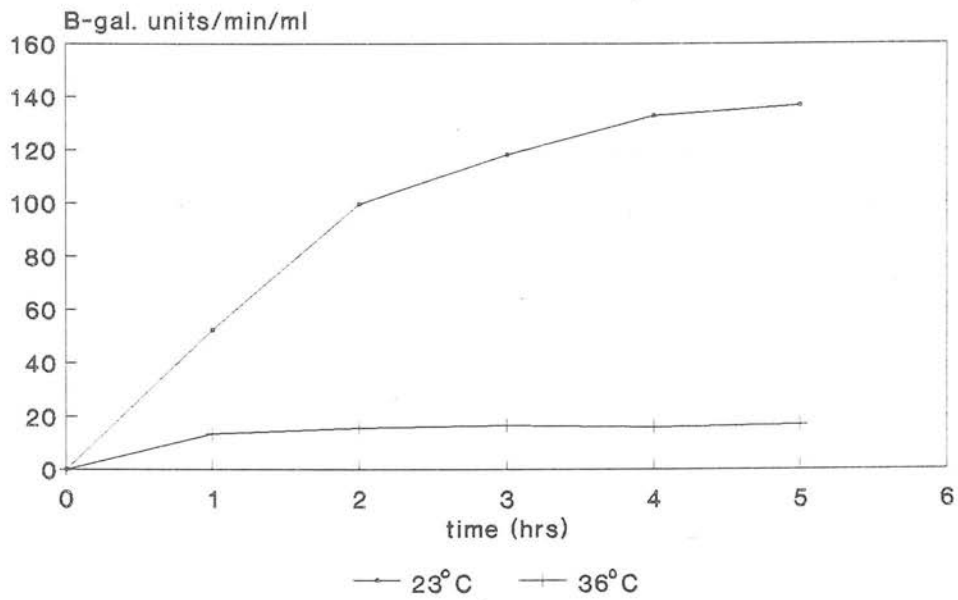
**FIGURE 4.6:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF SPJ8.31 (pJBM-5)  
INCUBATED AT 23°C AND 36°C; A *prp8* STRAIN ANALYSIS**

Cultures were grown at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 prior to galactose induction and heat-shock (see Chapter 2.2.9). Culture-samples (0.5ml) were taken at hourly intervals after galactose induction, and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately after the addition of galactose, the so-called time-zero (Chapter 2.2.9.1), and therefore 0 hours induction has an activity value of 0 (see Fig. 4.6).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

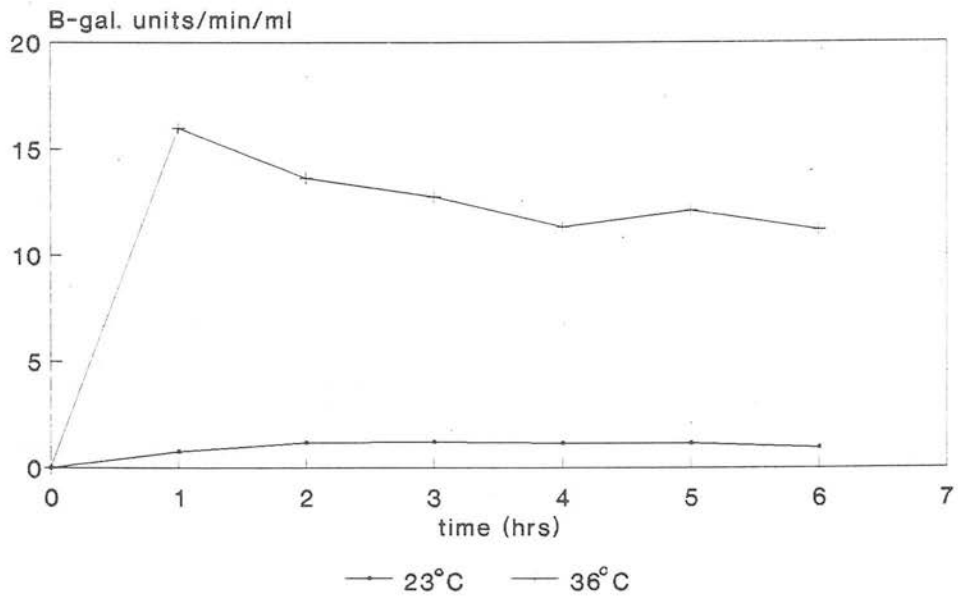
**FIGURE 4.7:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF SPJ8.31 (pJBM-4)  
INCUBATED AT 23°C AND 36°C; A *prp8* STRAIN ANALYSIS**

All experimental details were as described for SPJ8.31 (pJBM-5) in Fig. 4.6, except 0.5-2ml culture-samples were taken.

**FIG. 4.6**  
SPJ8.31 (pJBM-5)



**FIG. 4.7**  
SPJ8.31 (pJBM-4)



#### 4.1.2.4 DJY72 (*prp1*)

*Prp1* mutant strains are not thought to be defective in pre-mRNA splicing; however they are thought to be defective in the transport of RNA from the nucleus to the cytoplasm (for a fuller discussion see Chapter 5.6.3). It was decided to investigate the  $\beta$ -galactosidase profile of each plasmid when present in the *t<sup>s</sup> prp1* strain DJY72.

##### (a) pJBM-5

As with other strains a linear increase in activity was observed at 23°C (see Figure 4.8). However no increase in  $\beta$ -galactosidase was detected at 36°C i.e. there was no increase in background levels. This background measurement (activity prior to galactose induction) decreased with time, and was probably due to protein degradation as a consequence of cell death. This was further supported by the relatively quick arrest of cell division, as measured by the cellular optical density at 600<sub>nm</sub>, compared to the other *prp* strains at the non-permissive temperature (results not presented). This result could be interpreted as a complete block in mRNA transport at 36°C; however it could be that this strain was *t<sup>s</sup>* for galactose induction and/or that the *prp1* defect acted almost instantaneously to kill the cells. In both cases these mutations would prevent the detection of any increase in  $\beta$ -galactosidase activity. None of the strains which were *t<sup>s</sup>* for galactose induction had such a tight phenotype and they were only impaired in their ability to induce at the higher temperature; in DBY746 and SPJ0.9 this was roughly a 10-fold reduction (results not presented). These results suggest that at 36°C a *prp1-1* mutant is defective in the transport of mRNA from the nucleus to the cytoplasm. Other explanations are possible, such as a decrease in the stability of the RNA and/or protein at the elevated temperature.

(b) pJBM-4

To investigate the possibility that in a *prp1* mutant pre-and mRNA transport were separately controlled, the  $\beta$ -galactosidase activity profile of DJY72 (pJBM-4) was examined. As can be seen from Figure 4.9, activity was only detected when the DJY72 (pJBM-4) cultures were incubated at 23°C, and was completely absent from those incubated at 36°C. Thus if the *prp1* mutation affects RNA transport it acts on both pre-mRNA and mRNA. Cultures of DJY72 (pJBM-4) produced  $\beta$ -galactosidase activity which was 2.7% of that measured from those of DJY72 (pJBM-5) (incorporating the 3 to 5 hour time-points at 23°C and averaging their values), which was consistent with results from the other strains (see above).

(c) pJBM-4/s

In induction experiments utilising cultures of DJY72 (pJBM-4/s), only one time-point was assayed (5 hours at 23°C and 36°C results not presented). pJBM-4/s accounted for roughly 10% of the activity of pJBM-4 when the cultures were incubated at 23°C which was higher than those observed for DJY36 and SPJ8.31 but still considerably less than that of DBY745 (see Sections 4.1.2.1 to 4.1.2.3). Activity was not detected at 36°C, as would be expected of a defect in RNA transport.

It must be made clear that because sub-cellular fractionation of pre-mRNA to mRNA was not performed, it was not possible to determine directly that the *prp1* mutation inhibited the transport of RNA from the nucleus to the cytoplasm. However the results presented here are consistent with such a proposal.

#### 4.1.3 Summary of Results Indicating Screen Feasibility

The fusion-transcript pre-mRNA generated from the screening construct, pJBM-4, is spliced correctly and accumulates when splicing is inhibited. This accumulation can be detected

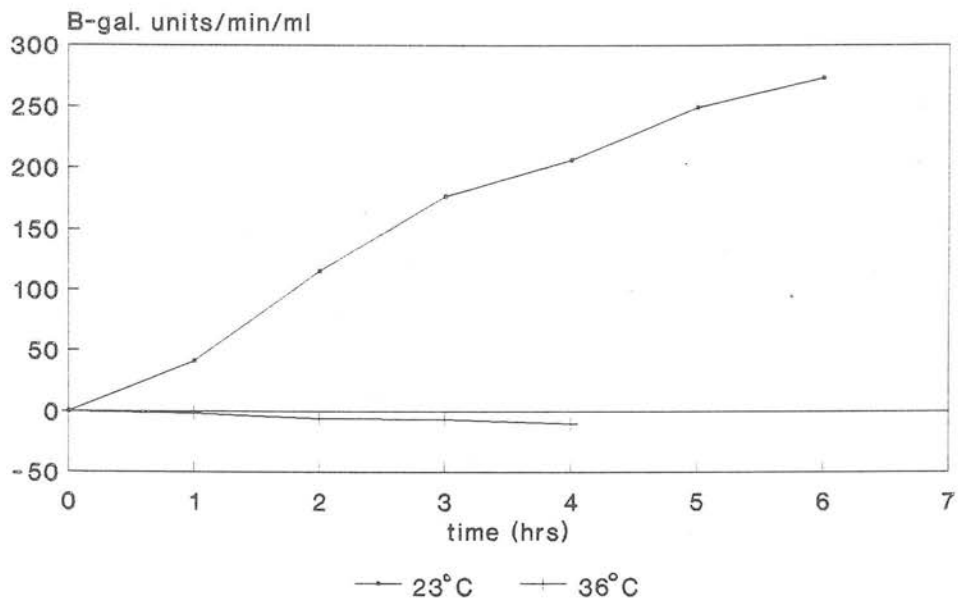
**FIGURE 4.8:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DJY72 (pJBM-5)  
INCUBATED AT 23°C AND 36°C; A *prp1* STRAIN ANALYSIS**

Cultures were grown at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 prior to galactose induction and heat-shock (see Chapter 2.2.9). Culture-samples (0.5ml) were taken at hourly intervals after galactose induction, and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately prior to the addition of galactose, the so-called time-zero (see Chapter 2.2.9.1), and therefore 0 hours induction has an activity value of 0 (see Fig. 4.8).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

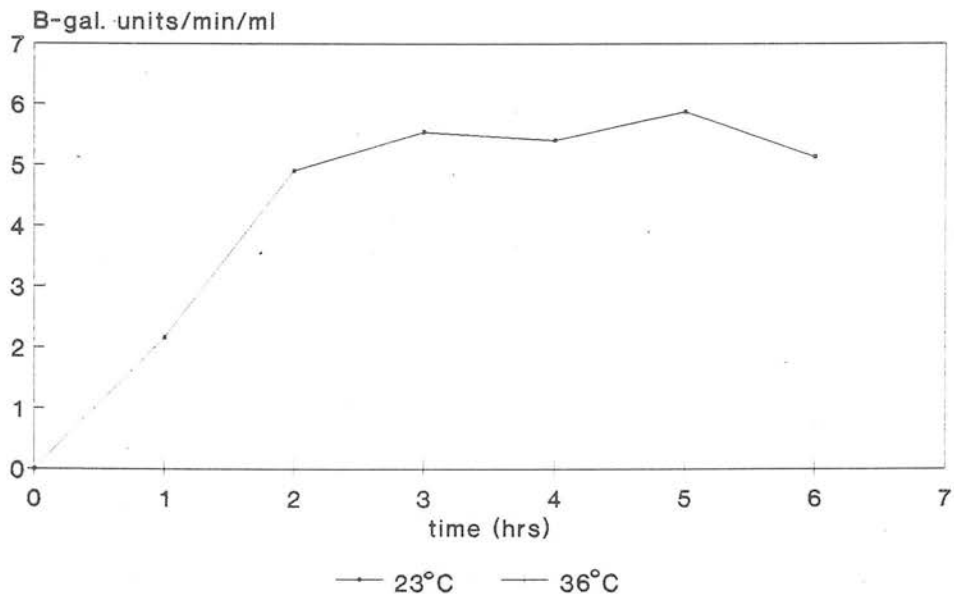
**FIGURE 4.9:  $\beta$ -GALACTOSIDASE ACTIVITY PROFILE OF DJY72 (pJBM-4)  
INCUBATED AT 23°C AND 36°C; A *prp1* STRAIN ANALYSIS**

All experimental details were as described for DJY72 (pJBM-5) in Fig. 4.8, except 0.5-2ml culture-samples were taken.

**FIG. 4.8**  
DJY72 (pJBM-5)



**FIG. 4.9**  
DJY72 (pJBM-4)



as an increase in  $\beta$ -galactosidase activity measured in *prp* strains incubated at the non-permissive temperature. Individual *prp* strains may produce different relative increases which could be a consequence of the stage in splicing at which they function. These results indicated that it was possible to set up a screening system using the rationale previously described (see Chapter 3.2).

## 4.2 DEVELOPMENT OF THE SCREENING PROCEDURE

From section 4.1.3 it can be seen that *prp* mutants harbouring the screening plasmid, pJBM-4, exhibit elevated levels of  $\beta$ -galactosidase activity when incubated at the non-permissive temperature. Obviously screening individual *r<sup>s</sup>* mutants for splicing defects by liquid culture assay would have been time-consuming and would therefore have defeated the point of developing a new screening procedure. Thus the results gleaned from the experiments described in sections 4.1.1 and 4.1.2 had to be transposed into a mass plate-screening procedure to accelerate detection of *prp* mutants. In the course of this transposition it became apparent that certain criteria had to be fulfilled to enable the screen to function. These points will be dealt with individually in the following sections; however it must be noted that they were intrinsically linked and relied on each other for the screen to operate efficiently. Figure 4.10 describes the steps involved in the screening procedure, from the creation of the *r<sup>s</sup>* mutant pool to the detection of potential *prp* mutants.

### 4.2.1 Choice of Wild-Type Strain

#### 4.2.1.1 Strains Producing "Petite" Mutants

On exposure to ultraviolet light some strains e.g. SI50-2B, produce a high percentage of "petite" mutants. This greatly hinders the isolation of *r<sup>s</sup>* colonies from the mutagenised pool.



#### 4.2.1.2 Glucose repressibility

As a consequence of the screening procedure adopted, all strains harbouring plasmids with the *GAL1* promoter had to be repressible for galactose induction when grown in the presence of glucose. It was also required that raffinose support growth of the chosen strain without affecting the induction of the *GAL1* promoter with galactose.

#### 4.2.1.3 Galactose induction mutants

As previously stated in section 4.1.2.4, some strains were  $t^s$  for galactose induction. At elevated temperatures it was possible that insufficient fusion-transcript was generated to produce detectable  $\beta$ -galactosidase activity. Indeed galactose induction was not possible in some strains (results not presented). A partial  $t^s$  phenotype (for galactose induction) did not appear to affect the screen (see Chapter 5.1); however the possibility exists that it determines the class of mutants that is recovered i.e. only those that show a large decrease in splicing efficiency which causes a large accumulation of pre-mRNA and consequently more  $\beta$ -galactosidase activity. The use of strains which are  $t^s$  for galactose induction may preclude the isolation of mutants with weak splicing defects because their small accumulation of pre-mRNA (splicing is relatively efficient) is further attenuated by the  $t^s$  galactose induction mutation which causes the production of less pre-mRNA transcript. In this respect it is relevant to note that the plate assay was essentially crude, and therefore  $\beta$ -galactosidase activity arising as a consequence of a *prp* mutation may have been sufficiently reduced as to prohibit the differentiation of a wild-type strain from a weak *prp* mutant.

There was no strain which fully satisfied all conditions; however both DBY746 (which was partially  $t^s$  for galactose induction) and DBY745 (which gave a high percentage of alternative splicing and/or translational read-through) were eventually utilised in the screening procedure.

#### 4.2.2 Growth Media

A variety of growth media was examined in various combinations e.g. sucrose, raffinose, glucose and galactose (results not presented). Some strains were not repressed when grown on sucrose e.g. SI50-2B. Galactose induction of the *GAL1* promoter was not a problem for most strains; however they could not utilise galactose efficiently as a carbon source and therefore raffinose (which was a non-inducing and non-repressing carbon source) was included. The inclusion of raffinose alleviated metabolic stress caused by carbon limitation which may have affected the assay. In the final method yeast cells were grown on glucose (YMG/Cas plates) and then transferred to galactose (YMR/Gal/Cas, see Chapter 2.1.2.2). This enabled full repression and induction of the yeast strains grown on these media.

#### 4.2.3 Detection of $\beta$ -galactosidase Activity

Initially yeast cells were grown on plates containing X-gal which was used to detect any production of  $\beta$ -galactosidase. These plates had to be buffered at pH7 to allow the  $\beta$ -galactosidase to function. Yeast cells require a pH of 6.5 for optimal growth and as a consequence grow extremely slowly under these conditions (other differences in media constituents may also have contributed to impair growth). The method finally utilised was to spot the cells onto nitrocellulose membranes placed on top of the relevant media. The cells were assayed by fracturing them in liquid nitrogen (while still on the membrane) and then placing the membrane into a petri-dish containing Z buffer and X-gal. Incubation of the membranes at 30°C identified  $\beta$ -galactosidase-producing colonies by virtue of their blue colour. For comparison of the activities generated by pJBM-5, 4 and 4/s in SI50-2B using this assay, see Fig. 4.11. Please note the difference in colourimetric intensity between these plasmids, especially pJBM-4 and 4/s. When residing in DBY745, pJBM-4 and pJBM-4/s did not exhibit such a contrast under various sensitivity levels tested (results not presented). This observation is consistent with the results indicating that the background activity exhibited by cultures of

DBY745 (pJBM-4/s) is greater than that of other strains (refer to Section 4.1.2). Spotting the cells onto nitrocellulose membranes allowed greater flexibility in the design of the screening procedure by enabling easy transfer of colonies between different media without changing the number of cells present in each colony (a problem inherent in replica plating). Because of the background level of activity (see Section 4.1.2.1), it was important that on individual plates the colonies grew at roughly the same rate and therefore contained a similar number of cells. This minimised the variation in background activity produced as larger colonies produced greater total amounts of background  $\beta$ -galactosidase activity due to their increased cell number. Thus slow growing mutants were assayed separately. Fig. 4.12 shows an example of duplicate membranes containing a number of  $t^s$  mutants which were screened using the plate-assay. One of the mutants in this screen clearly exhibits a blue colour when incubated at 36°C only (marked by the arrows), while the wild-type control (pJBM-4-containing colony) is still colourless. Therefore this screen identified a possible *prp* mutant. Cells containing pJBM-5 were used as a positive control to detect efficient induction of the *GAL1* promoter.

Wild-type cells produce a background level of activity caused by the leakage of pre-mRNA from the nucleus (see Section 4.1.2). In order to detect any increase in activity (taking into account the crude nature of the plate assay), the sensitivity of detection had to be reduced below the leakage activity. Thus normal wild-type (for splicing) cells would remain colourless even though pre-mRNA could exit the nucleus and produce  $\beta$ -galactosidase activity (this "leakage" activity was detected by increasing the sensitivity of the assay; results not presented). The sensitivity of the assay was regulated by including different amounts of substrate (X-gal) in a constant volume of Z buffer (see Fig. 4.10). However other methods could have been utilised e.g. stoppage of the assay reaction in a time-dependent manner, limiting the time of galactose induction and/or reducing the time of growth prior to induction. It was found that the most effective and convenient method of *prp* mutant detection was to grow the  $t^s$  mutants at 23°C on a *GAL1* repressing medium, transfer them to an inducing medium and incubate duplicate samples of the cells at both 23°C and 36°C, and then assay them thereafter at hourly intervals. Due to the increase in cell number the wild-type cells eventually produced enough

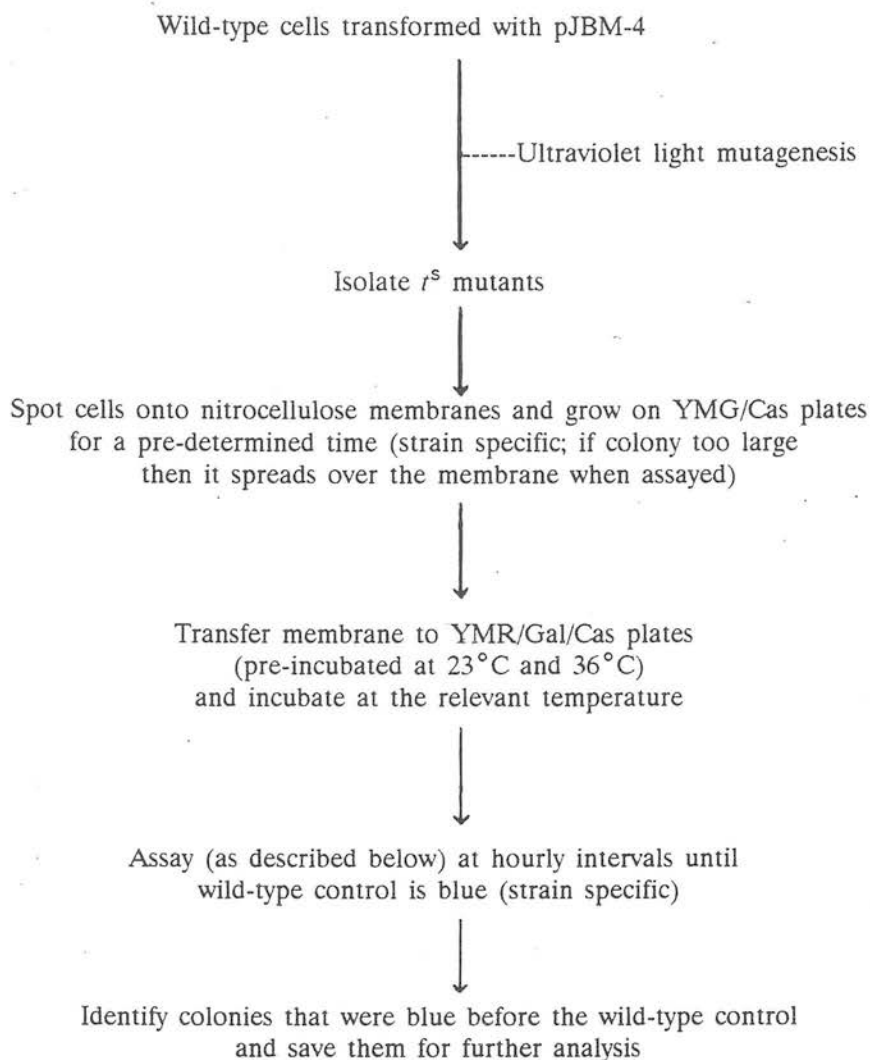
$\beta$ -galactosidase activity to turn blue. However this increase in  $\beta$ -galactosidase activity could also be a consequence of the continual induction of the *GAL1* promoter, which may cause the titration of factors necessary for the commitment of pre-mRNA transcripts to splicing, and therefore lead to the exit of more pre-mRNA transcripts from the nucleus. The period before the wild-type cells went blue was a "window" in which increased  $\beta$ -galactosidase activity (hence pre-mRNA) could be detected, and the time-scale of this "window" depended on the sensitivity employed in the assay.

#### FIGURE 4.10: THE SCREENING PROCEDURE

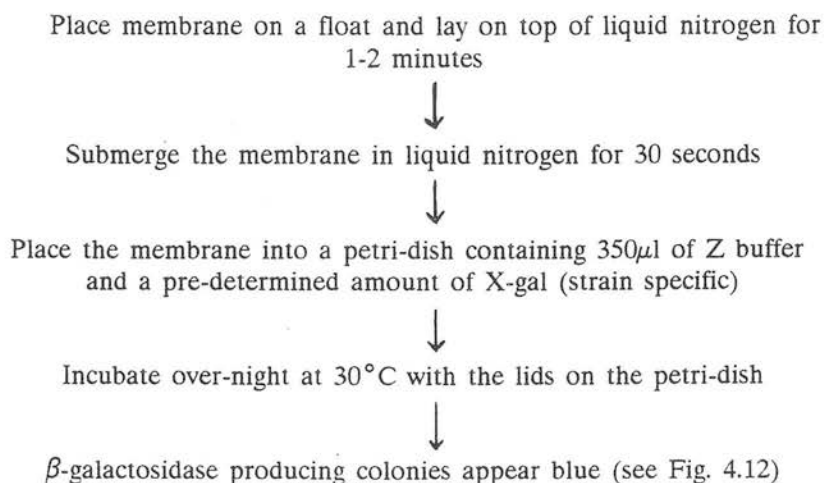
The screening procedure is given, from the creation of a bank of  $t^s$  mutants to the final plate-assay method utilised to identify possible *prp* mutants.

FIG. 4.10

# SCREENING PROCEDURE



## Assay Technique



**FIGURE 4.11: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES ARISING FROM pJBM-5, -4 AND -4s IN S150-2B, USING THE PLATE-ASSAY PROCEDURE**

S150-2B transformants containing the relevant plasmids were suspended in water and spotted onto a nitrocellulose membrane placed on top of YMG/Cas medium (see Chapter 2.1.2.2). The plate was incubated at 23°C for 21 hours and then the membrane was transferred to YMR/Gal/Cas medium and incubated at 23°C for a further 5 hours. After placing on a float and leaving on liquid nitrogen for 1-2 minutes, the membrane was submerged for 30 seconds and then placed into a petri-dish containing 350 $\mu$ l of Z buffer and 10 $\mu$ l of 20mg/ml X-gal. It was then incubated at 30°C overnight. The  $\beta$ -galactosidase activity arising from the plasmids is clearly visible as a blue colour and the intensities of colour correlate with each plasmids expected  $\beta$ -galactosidase profile.

# FIG. 4.11

SI50-2B

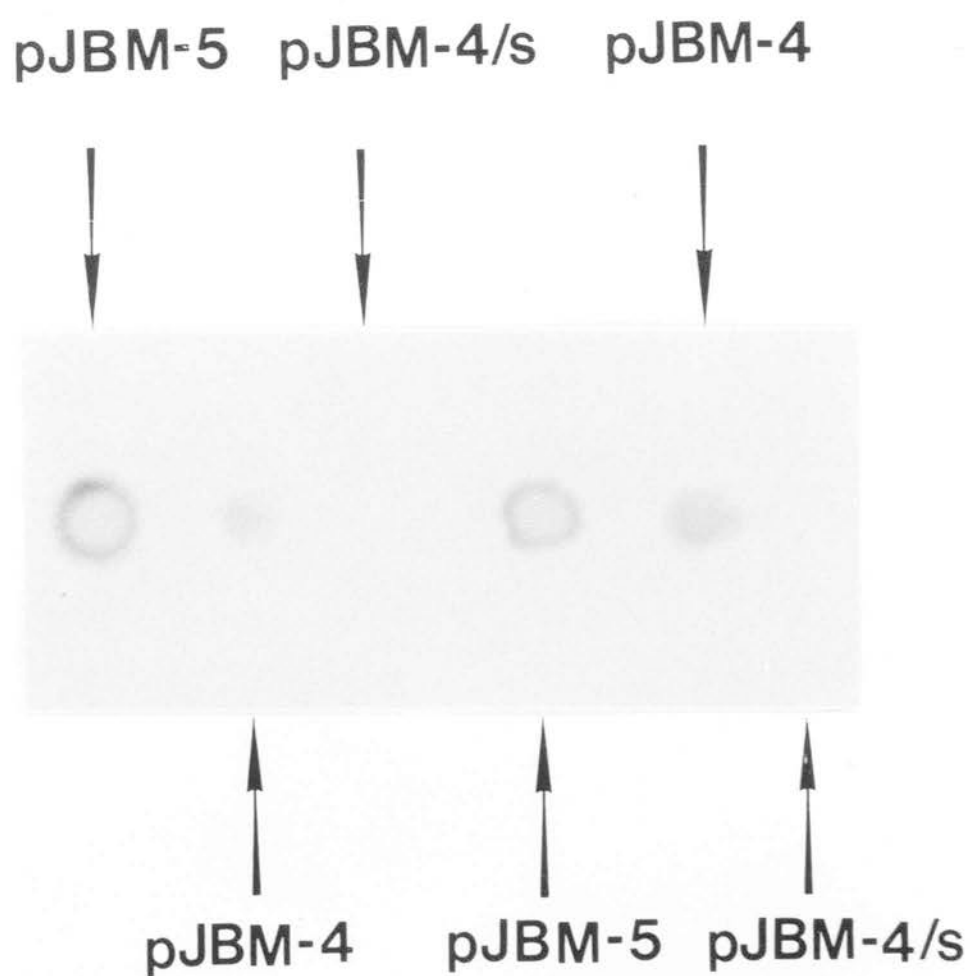


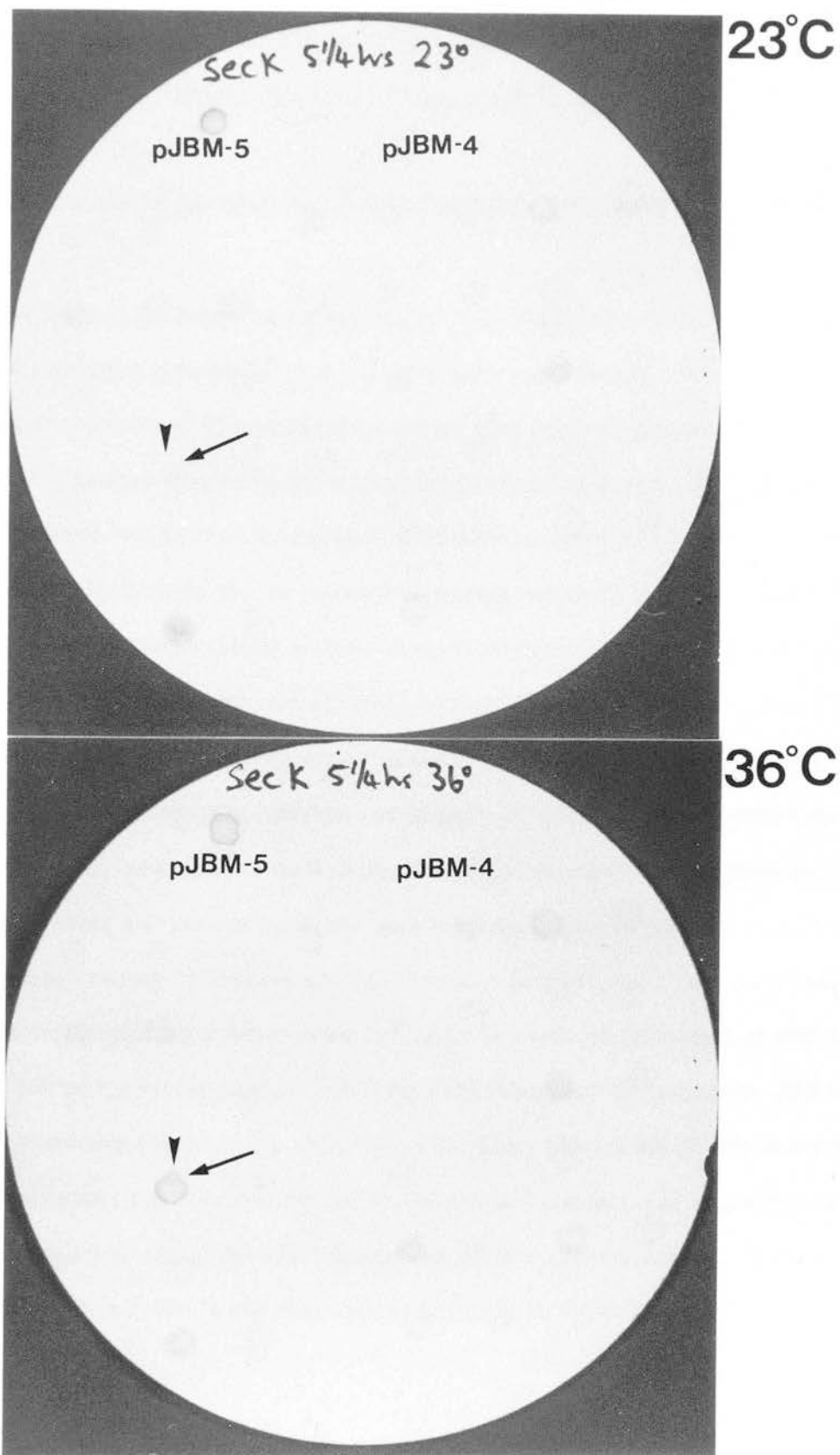


FIGURE 4.12: IDENTIFICATION OF A POTENTIAL *prp* MUTANT USING THE PLATE-SCREEN ASSAY

A number of *r<sup>s</sup>* mutants were screened for a potential *prp* phenotype. The *r<sup>s</sup>* mutants were suspended in water and spotted, in the same grid pattern, onto two nitrocellulose membranes which had been placed on top of two plates of YMG/Cas medium. Each plate was then incubated at 23°C for 21 hours. The two membranes were then placed onto YMR/Gal/Cas plates which had been pre-incubated at 23°C and 36°C, and thereafter incubated at the relevant temperature for five and a half hours. Thus identical *r<sup>s</sup>* colonies, placed on duplicate membranes, were induced at 23°C and 36°C. One colony exhibited a blue colour (and hence increased levels of  $\beta$ -galactosidase activity) after incubation at 36°C only (in fact it is 60h; see Chapter 5.3.1) and is thus identified as a potential *prp* mutant (see the arrows).

The wild-type control transformants on the membranes presented in this figure clearly indicate that there is induction of the *GAL1* promoter at both temperatures (pJBM-5), and that the screening plasmid (pJBM-4) does not produce any detectable activity at any temperature under these assay conditions. The relevant control transformants have the name of the plasmids they contain indicated below them.

FIG.4.12



## CHAPTER 5

### ISOLATION AND CHARACTERISATION OF POTENTIAL CONDITIONAL-LETHAL *prp* MUTANTS

#### 5.1 GENERATION OF TEMPERATURE-SENSITIVE MUTANTS

As indicated in Chapter 1.9, the existence of conditional-lethal mutants provides several advantages when studying molecular processes. In order to isolate conditional-lethal *prp* mutants a bank of  $t^s$  mutants was created, using ultraviolet light (UV) to generate this pool (see Chapter 2.2.3.7). Apart from being the method in use in this laboratory UV irradiation has several advantages over other procedures such as chemical mutagenesis; it is easily controllable and relatively safe. It might also be expected to produce mutations that differ from those previously detected by other screens in which chemical mutagens were used to generate the  $t^s$  pool (Hartwell, 1967; Vijayraghavan *et al.*, 1989). Two wild-type strains, DBY745 and DBY746, were chosen for mutagenesis. In generating  $t^s$  mutants a balance must be obtained between the amount of UV energy expended (duration and intensity of exposure), and the percentage of mutants recovered. Long exposure leads to the introduction of multiple  $t^s$  mutations into the genetic background, and this can hinder the analysis of the relevant  $t^s$  gene and the isolation of its wild-type counterpart. Because of this, and also because each strain has a distinct mutational repair response, a separate survival curve, in which the percentage of cells that survive is plotted against the duration of UV exposure, is required for each strain (this also applies to chemical mutagens; Miller, 1972; Haynes and Kunz, 1981). It was considered that the original *prp* mutants had been over-exposed to the chemical mutagen used to generate them (chemical mutagenesis giving 99%-99.8% killing) and this had often resulted in the presence of more than one  $t^s$  mutation in any given mutant strain e.g. *prp6*, *prp7* and *prp9*. A value of

80%-90% killing is more likely to produce a single  $r^s$  mutation per genome (Maddock, J., pers. comm.). In a recent screen 70%-90% killing was the value utilised (Vijayraghavan *et al.*, 1989).

Survival curves were generated for DBY745 and DBY746, and the appropriate exposure time, calculated to produce 80%-90% cell death (~1 minute for DBY746 and ~1 minute 15 seconds for DBY745) was utilised to create the  $r^s$  pool from each strain (see Fig. 5.8 for the survival curves generated for each strain). In total 590  $r^s$  mutants representing 307 independent isolates, i.e. mutants isolated from different cell pools, were generated (296  $r^s$  mutants in total from DBY745, 129 of which were independent isolates, and 294  $r^s$  mutants from DBY746 of which 177 were independent isolates) and tested for pre-mRNA processing defects in the screening procedure as described in Chapter 4.2. It should be noted that because the screening procedure is relatively fast the discrimination of a  $r^s$  phenotype was not conclusive, thus not all 590 mutants were in fact  $ts$ , although the vast majority were. This needs to be taken into account when discussing the efficiency of the screen in comparison to others (see the Discussion at the end of this Chapter).

## 5.2 SCREENING FOR POTENTIAL *prp* MUTANTS

All 590  $r^s$  mutants were screened for their ability to produce elevated levels of  $\beta$ -galactosidase activity as detailed in Chapter 4.2.3, Fig. 4.10. However the strains used to generate the  $r^s$  mutants respond differently to galactose induction and thus the conditions of assay are unique to each one. In reference to Fig. 4.10, DBY745 requires 21 hours growth on YMG/Cas and 10 $\mu$ l of 20mg/ml X-gal in the assay, compared to 41 hours and 25 $\mu$ l for DBY746. Because the screen was being tested for the first time, colonies producing a relatively light blue phenotype were also isolated for further analysis.

## 5.2.1 DBY746

Out of 294  $t^s$  mutants 11 colonies were isolated for further analysis. The blue phenotype exhibited by the individual colonies at the two temperatures is given in Table 5.1. Please note that the intensity of colour varied between colonies but is not indicated.

TABLE 5.1 DBY746 DERIVED  $t^s$  MUTANTS EXHIBITING A BLUE PHENOTYPE

<u>Colony</u>	<u>Blue Phenotype</u>		<u>Colony</u>	<u>Blue Phenotype</u>	
	<u>23°C</u>	<u>36°C</u>		<u>23°C</u>	<u>36°C</u>
82a	+	-	235f	+	-
193d	+	-	254	+	-
196a*	?	?	275	-	+
221c	+	+	298	+	-
234b	+	-	316b	+	-
196b*	?	?			

((+) or (-) indicates the presence or absence respectively of a blue phenotype. (?) indicates that the temperature distinction in blue colour was not noted. \* indicates mutants isolated from the same aliquot).

## 5.2.2 DBY745

Thirty-five colonies were isolated for further analysis out of the 296  $t^s$  mutants originally generated from this strain. The blue phenotype exhibited by these colonies at the two temperatures is given in Table 5.2. Please note that the intensity of colour varied between the colonies but is not indicated in this analysis.

TABLE 5.2 DBY745 DERIVED *t<sup>s</sup>* MUTANTS EXHIBITING A BLUE PHENOTYPE

<u>Colony</u>	<u>Blue Phenotype</u>		<u>Colony</u>	<u>Blue Phenotype</u>	
	<u>23°C</u>	<u>36°C</u>		<u>23°C</u>	<u>36°C</u>
4c	+	+	63d	+	-
6c*	+	-	70a	+	+
6g*	+	-	83a	+	-
7c	+	+	101g*	+	-
9c	+	+	102e	-	+
19a	+	+	104b	-	+
20a	+	-	105c	-	+
24a*	+	+	109i	+	+
31a	+	-	110a	+	-
34a	+	+	114e	-	+
36a	+	-	121e*	-	+
43c	+	+	124c	-	+
51a	+	-	139b	+	-
52b	+	-	140c	+	-
60h	-	+	146d	+	-
61d	+	+	148c	+	-
24b*	+	-	101h*	+	+
121a*	+	-			

((+) and/or (-) indicates the presence or absence respectively of a blue phenotype : \* indicates mutants isolated from the same plate; compare the colony numbers to identify relevant pairs).

### 5.3 NORTHERN BLOT ANALYSIS OF POTENTIAL *prp* MUTANTS

RNA preparations were made from the 46 potential *prp* mutants after their incubation at 23°C and 36°C (see Sections 5.2.1 and 5.2.2 for the mutants used in this analysis). The heat-shock at 36°C usually lasted 1.5-2 hours, in which time it was thought that any splicing defect would become apparent. This supposition is based on the pre-mRNA accumulation profiles of *prp2* to *prp11*; however this general time-scale of heat-shock and pre-mRNA accumulation may not allow detection of all types of *prp* mutants, other certain defects may only be detected on longer incubation at the elevated temperature (see the Discussion at the end of this Chapter). Three genes were utilised as probes to detect the *prp* phenotype in the Northern blot experiments.

(a) Actin; the actin pre-mRNA transcript is spliced efficiently and served as a comparison with the *CYH2* pre-mRNA.

(b) *CYH2*; this gene encodes the ribosomal protein L29, and its transcript is spliced relatively inefficiently. This was used to help distinguish those mutants which had a less pronounced effect on splicing, as it was presumed that small decreases in splicing efficiency would produce a more noticeable effect on pre-mRNA to mRNA ratios when an inefficiently spliced transcript was studied.

(c) *GPM*; the phosphoglycerate mutase transcript has no intron and served as an internal control of RNA levels in each preparation. Thus it was possible to determine that any increase in pre-mRNA at the non-permissive temperature was not due to an increase in the general level of RNA.

The criterion used to identify a possible defect in pre-mRNA processing is that the ratio of mRNA : pre-mRNA decreases at 36°C. This is based on a proposed steady-state kinetic

model for the generation of spliced mRNA (Pikielny and Rosbash, 1985). Briefly, two possible scenarios are considered to explain the phenotype of mutations which affect splicing;

(a) there are mutations that reduce the efficiency of splicing below the rate of turnover of the pre-mRNA and mRNA. Because splicing is drastically inhibited, mRNA is not generated while the existing mRNA is degraded in the cell. This results in rapid mRNA depletion. Although degraded relatively quickly (compared to mRNA), pre-mRNA accumulates in such mutants because it is not converted to mRNA. Both RNAs can therefore reach a level which is dependent on the splicing efficiency of the pre-mRNA and the turnover rates of the pre-mRNA and mRNA. This phenotype is characteristic of the *prp2* to *prp11* mutants.

(b) There are mutations which reduce the efficiency of splicing but not below the turnover rates of the RNAs. Because mRNA is degraded relatively slowly (therefore relative to pre-mRNA there is more of it present at any one time) the small decrease in splicing efficiency does not discernably affect the mRNA pool size. In this way the mRNA levels appear to remain constant. In contrast, the pre-mRNA is degraded relatively quickly (therefore at any one time there is less of it about); thus small changes in splicing efficiency have a much more pronounced effect (relatively) on the pre-mRNA pool size. It should be noted that mutants showing this phenotype could have a *t<sup>s</sup>* defect in a RNase activity which results the pre-mRNA accumulation observed at the non-permissive temperature.

The consequence of these two scenarios is that the pre-mRNA level observed at 36°C need not be greater than that monitored at 23°C only that the ratio of pre-mRNA to mRNA increases. Another consequence is that in pre-mRNA splicing mutants the mRNA level need not change; however in these cases the pre-mRNA level should increase.

Both scenarios can be described as a relative accumulation of pre-mRNA. On the basis of the results obtained from Northern blot analysis the *t<sup>s</sup>* mutants isolated from the screen were placed into 4 groups.



### 5.3.1 *t<sup>s</sup>* Mutants Showing a *prp* Northern Phenotype

*Prp2* to *prp11* mutants accumulate pre-mRNA at the expense of mRNA when incubated at 36°C. This pattern is characteristic of such "classic" *prp* mutants and several of this type were identified in this screen. Fig. 5.1A shows the Northern phenotype exemplified by the two wild-type strains used in the screen (DBY745 and DBY746) and a *prp8* strain (SPJ8.31). Fig. 5.1B shows the Northern blot phenotype exhibited by the *t<sup>s</sup>* mutants 221c, 275, 298, 9c and 104b, and Fig. 5.2 those of 105c, 109i, 60h, 61d and 124c. All mutants presented in Figs. 5.1B and 5.2 were isolated utilising the screening procedure described in Chapter 4, Fig. 4.10. It should be noted that compared to the others 60h, 61d and 109i do not produce such a marked increase in pre-mRNA and indeed could also be included in section 5.3.2. In the analysis of 60h and 61d it is apparent that the RNA preparations made from these strains (incubated at 23 C only) are defective (see tracks 5 and 7). The reason for this is unknown, however other RNA preparations made from the same strains do not exhibit this defect (results not presented). The Northern blot of 124c (see Fig. 5.2, tracks 9 and 10) is over-exposed, however it is apparent that it accumulates pre-mRNA at the expense of mRNA when incubated at 36°C though there is still a considerable level of mRNA present. Though it appears to lose more mRNA when incubated at 36°C, 124c produces a Northern blot phenotype similar to 121e and 196a (see Section 5.3.2) and could be included in that section. Thus 10 mutants demonstrating a *prp* Northern phenotype have been identified.

### 5.3.2 Mutants Which May Have Been Defective in Splicing

There were 5 mutants in this class (20a, 36a, 109i, 121e and 196a), and each will be dealt with separately.

## FIGURE 5.1: NORTHERN BLOT HYBRIDISATION ANALYSIS OF POTENTIAL *prp* MUTANTS

RNA was prepared from yeast strains which had been grown in YPDA at 23°C or shifted to 36°C for 1.5-2 hours (temperature indicated above each track). RNA (10-50µg) was denatured with glyoxal, electrophoresed through a 1.2% or 1.5% agarose gel and blotted onto a GeneScreen membrane (see Chapter 2.2.8). These were then hybridised with the uniformly labelled 1038bp *Bam*HI/*Bgl*II fragment of pYA301 (actin transcript probe), the 2kb *Eco*RI fragment from pBRCYH-3 (*CYH2* transcript probe) and the 1.3kb *Hind*III/*Sal*I fragment of pUC-GPM (phosphoglycerate mutase transcript probe) (see Chapter 2.1.5, Table 2.3). Only one probe was used at any one time, after which the membrane was stripped of radioactivity and rehybridised with another. DNA restriction size markers were usually run with each experiment, however they are not shown because the transcripts analysed are not presented in the order of size. The sizes of the relevant RNA species (as indicated in the figure) are as follows;

Actin - pre-mRNA - approx. 1700nts	<i>CYH2</i> - pre-mRNA - 1082nts
mRNA - approx. 1392nts	mRNA - 572nts

*GPM* - mRNA - approx. 1100nts

A. Northern blot analysis of DBY745, DBY746 and SPJ8.31. The results for these strains are presented to enable a comparison with the potential *prp* mutants. All of the newly isolated *t<sup>s</sup>* mutants were derived from DBY745 and DBY746 (see Section 5.2).

B. Northern blot analysis of potential *prp* mutants.

FIG. 5.1

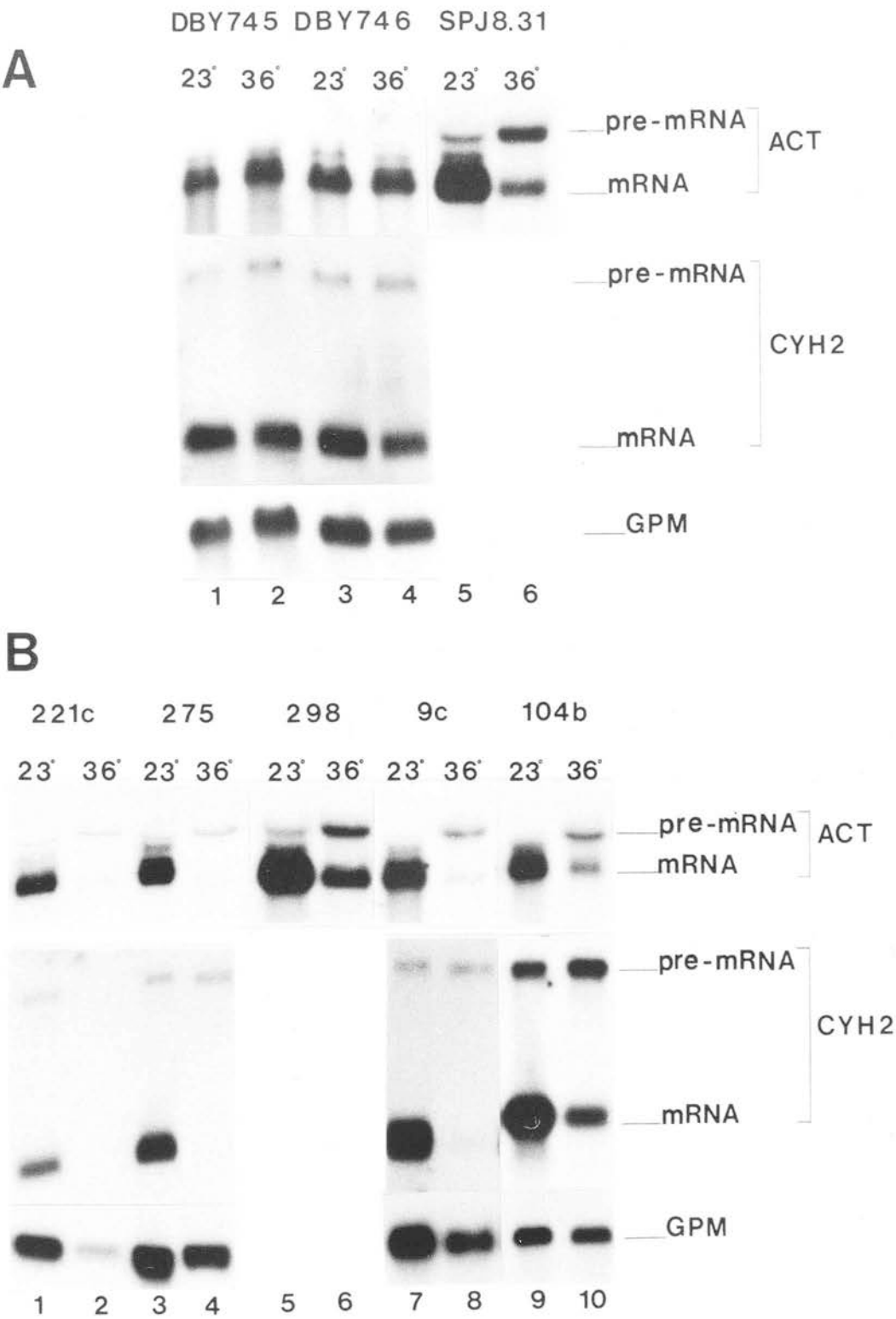
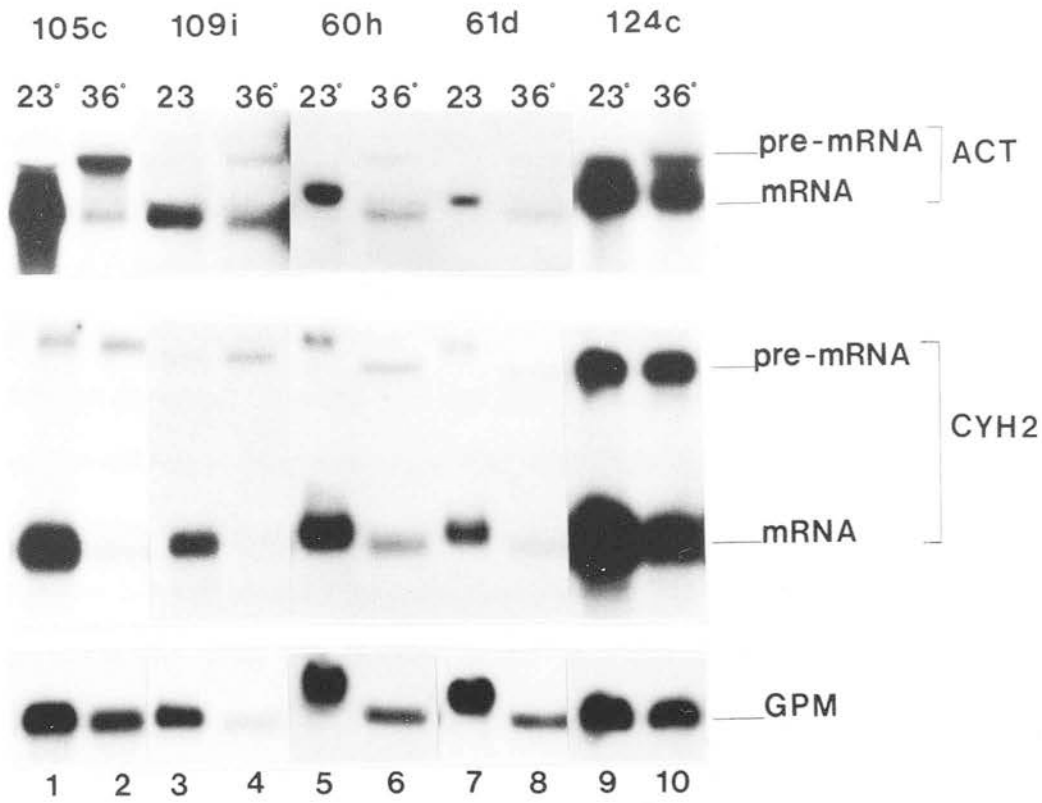


FIGURE 5.2: NORTHERN BLOT HYBRIDISATION ANALYSIS OF POTENTIAL *prp* MUTANTS

All experimental details were as described for Fig. 5.1. The sizes of the relevant RNA species are as for Figure 5.1.

FIG. 5.2



A slight increase in the ratio of pre-mRNA to mRNA of the actin and *CYH2* pre-mRNA transcripts has been detected when this mutant is incubated at 36°C (see Fig. 5.3A and compare tracks 1 and 2). RNA preparations made at a later date (see below) did not show any accumulation of actin pre-mRNA whereas the *CYH2* pre-mRNA remained at a constant level even though its mRNA level decreased, and this could be interpreted as indicating a splicing defect (see Fig. 5.3A, tracks 3 and 4). To determine whether efficient accumulation of pre-mRNA depended on a longer duration of heat-shock, RNA was prepared at hourly intervals up to and including 4 hours from cells incubated at 36°C (see Fig. 5.3A, tracks 5 to 8; please note that the culture was in the logarithmic stage of growth when the heat-shock was applied, results not presented). However, actin and *CYH2* pre-mRNA increased in concert with their respective mRNAs at every time-point. The variation in the Northern phenotype was at first taken to mean that the mutant was unstable, as the *prp* phenotype appeared to disappear with consecutive RNA preparations (cells stored on plates and RNA preparations made in the order given in Fig 5.3A, tracks 1-4). However, recent indications suggest that detection of a splicing defect may require longer heat-shock and there is a possibility that it may be defective in snRNP assembly (see 36a below and the Discussion at the end of this chapter). This mutant is by far the slowest growing *r<sup>s</sup>* mutant isolated and grows better on YPDA than on YMG/Cas. Its genotype is as DBY745 i.e. *leu2*, *ade2-100* and *ura3-52* (results not presented), therefore the reason for its slow growth on YMG/Cas is not due to any auxotrophic mutation and remains unknown.

In the very first Northern blot performed using RNA isolated from this mutant, accumulation of the actin pre-mRNA with a concomitant decrease in mRNA was detected (see Fig. 5.3B, tracks 1 and 2); however this was the only gene transcript analysed. Subsequent

analysis using different RNA preparations revealed a similar situation to that encountered with 20a i.e. the *prp* phenotype disappeared with consecutive RNA preparations. It is interesting to note that *CYH2* RNA is almost completely absent at 36°C even though the GPM control indicates that RNA depletion is only 3-4-fold (see Fig. 5.3B, tracks 3 to 6). The reason for the depletion may be due to the possibility that the RNA was isolated from cells which had reached the late logarithmic stage of growth before the heat-shock was applied, as the optical density of the culture used in this experiment was not directly monitored. To test this, and also to extend the duration of heat-shock, a similar time-course experiment to that carried out with 20a was performed (see Fig. 5.3B, tracks 7 to 10). Aliquots were removed at hourly time-points from a culture incubated at 36°C. The OD 600<sub>nm</sub> measurements each time-point were 0.217, 0.28, 0.31 and 0.336 respectively; therefore the cells were in the logarithmic phase of growth when the heat-shock was applied (please note that the same was true for all the time course experiments i.e. 20a, 109i and 121e also, results not presented). Results clearly indicate that for both actin and *CYH2* the RNA level decreases up to 3 hours and then increases to roughly the 1 hour level (this pattern is mirrored by GPM); however there is a substantial increase in *CYH2* pre-mRNA at the 4 hour time-point relative to the 1 hour (compare tracks 7 and 10: a slight accumulation of the actin pre-mRNA is also visible at the 4 hour time-point, see track 10). The reason for the increase in total RNA after 4 hours is not known and needs to be verified, though it could simply be the result of a more concentrated RNA preparation. It is possible that the initial decrease in RNA levels observed for 36a after incubation at 36°C is a general, transient heat-shock effect, however the wild-type strains, DBY745 and DBY746, do not exhibit a decrease in RNA levels after two hours incubation at 36°C (see Fig. 5.1). This suggests that the loss of RNA is a consequence of the *r<sup>s</sup>* mutation(s) (or other mutations) present in 36a. Detection of a blue phenotype in the plate assay when 36a was incubated at 36°C required 6 hours of galactose induction and heat-shock (results not presented) and could reflect this lag time in pre-mRNA accumulation (most mutants produced the blue colouration within 2-4 hours). These results suggest that on longer incubation 36a accumulates significant levels of pre-mRNA.

#### FIGURE 5.3A: NORTHERN BLOT HYBRIDISATION ANALYSIS OF MUTANT 20a

- Tracks 1-4: RNA was prepared from cultures of 20a grown in YPDA at 23°C or shifted to 36°C for 1.5-2 hours (see Chapter 2.2.8). All other experimental details and RNA transcript sizes were as for Fig. 5.1. Tracks 1 and 2 represent one experiment and tracks 3 and 4 a repeat experiment carried out at a later date.
- Tracks 5-8: Time-course of Northern blot phenotype exhibited by 20a on incubation at 36°C. Cultures of 20a were initially grown in YPDA at 23°C to an OD 600<sub>nm</sub> of 0.1-0.4 and then shifted to 36°C. Thereafter, RNA was prepared at hourly intervals (1-4 hours, tracks 5-8 respectively).

#### FIGURE 5.3B: NORTHERN BLOT HYBRIDISATION ANALYSIS OF MUTANT 36a

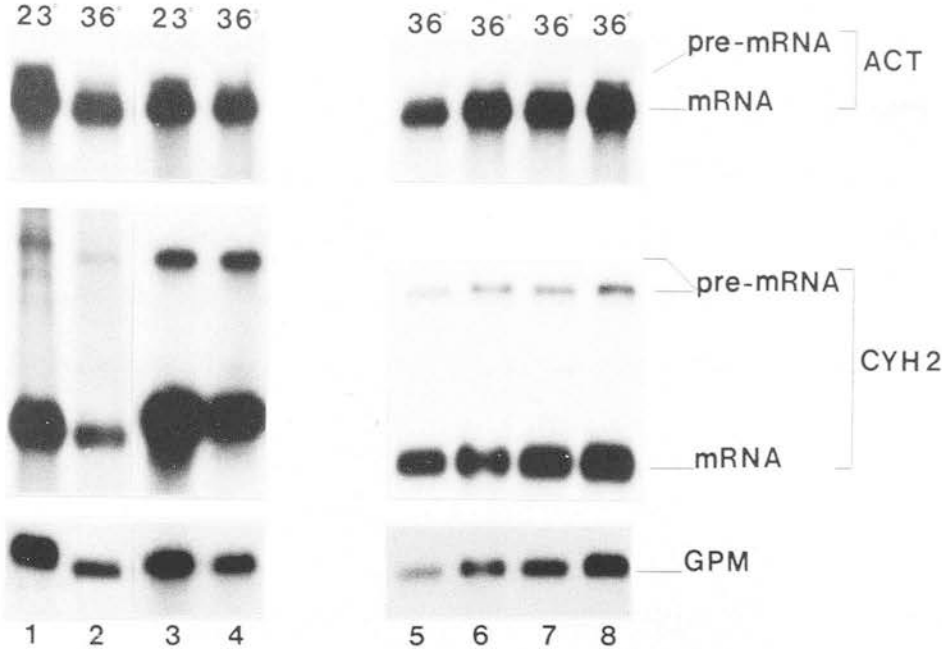
- Tracks 1-6: Experimental details as described for Fig. 5.3A. The *CYH2* or *GPM* transcripts were not analysed in the RNA preparations used in tracks 1 and 2. Tracks 1 and 2 represent one experiment and tracks 3 and 4, and tracks 5 and 6 represent two individual repeat experiments carried out at later dates.
- Tracks 7-10: Time-course of Northern blot phenotype exhibited by 36a on incubation at 36°C. Cultures of 36a were initially grown in YPDA at 23°C to an OD 600<sub>nm</sub> of 0.1-0.4 and then shifted to 36°C. Thereafter, RNA was prepared at hourly intervals (1-4 hours, tracks 7-10 respectively).



FIG.5.3

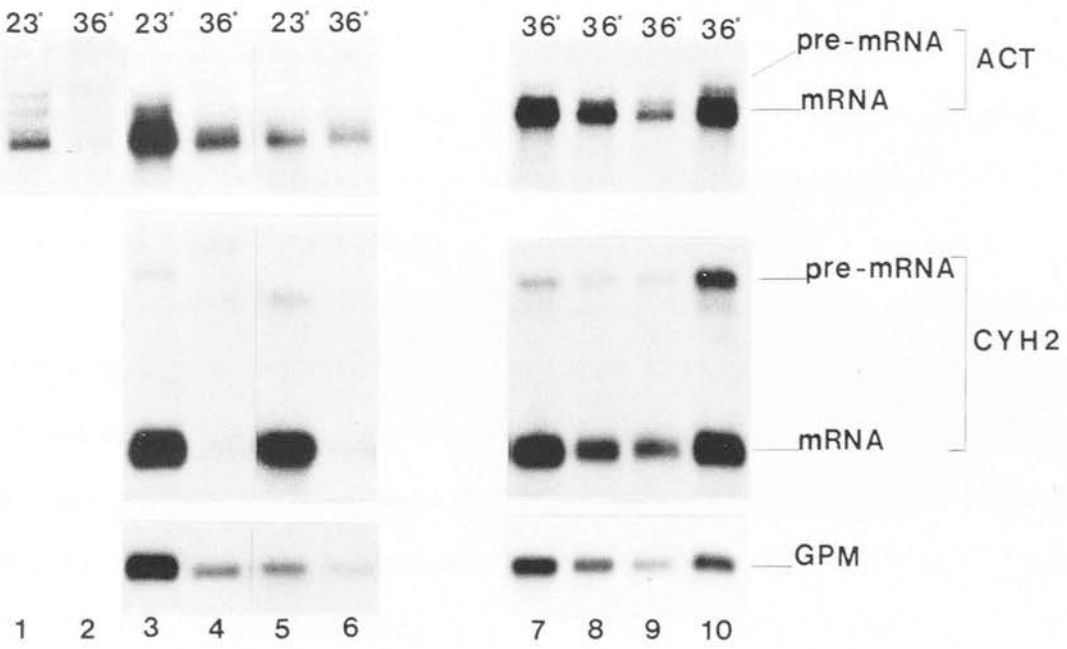
A

20a



B

36a



As indicated in 5.3.1, 109i shows a definite *prp* phenotype. However in time-course experiments (the same as those performed on 20a, 36a and 121e), pre-mRNA increases relative to a maintained mRNA level. This effect lasts for 3 hours, after which both the pre-mRNA and mRNA levels decrease (see Fig 5.4C). Such a result could indicate an increase in pre-mRNA stability with no defect in splicing. If the *r<sup>s</sup>* mutation is the cause of this observation then it must mean that accumulation of pre-mRNA is detrimental to the cell and/or that the mutation has pleiotropic effects because the cells cease to grow. The reason for the difference between this observation and those utilising different RNA preparations is not known, and indeed recent results repeat the *prp* phenotype seen in Fig. 5.2 (results not presented).

## 5.3.2.4 121e

This mutant appears to accumulate pre-mRNA when incubated at both the permissive and non-permissive temperatures; however there appears to be relatively more pre-mRNA present at 36°C than at 23°C (see Fig. 5.4A, tracks 1 and 2). When 121e was incubated at 36°C the decrease in mRNA level was not substantial and therefore it was considered that a longer heat-shock might be required to produce a significant contrast in the changes of pre-mRNA and mRNA levels. Fig. 5.4A, tracks 3 to 6 shows the results of a time-course experiment similar to those carried out for 20a and 36a (see above). Mutant 121e produces a Northern phenotype similar in outline to 109i (see Section 5.3.2.3) i.e. the pre-mRNA levels of both transcripts increase up to three hours and then decrease. At the same time the mRNA level appears to stay at a similar level for the three hours and then also decreases. However in contrast to 109i the classic *prp* phenotype has not been detected for 121e (compare Fig. 5.2, tracks 3 and 4 with Fig. 5.4A, tracks 1 and 2). It should be noted that comparison of 121e with 124c could mean that 124c should be included in this group (see Fig. 5.2), however 121e does not appear to lose so much mRNA at 36°C though the autoradiographs are over-exposed and caution must be

observed when interpreting these data. Mutant 124c cannot complement a *prp24* mutant and is therefore likely to be allelic to *prp24* (see Section 5.4), and indeed this is a reason for postulating that 121e is in fact a *prp* mutant defective in splicing and also for placing 124c in Section 5.3.1.

#### 5.3.2.5 196a

Mutant 196a exhibits a Northern phenotype similar to 121e (see Fig. 5.4B) though in 196a there appears to be more pre-mRNA accumulation at 23°C. Mutant 196a is in fact a *prp8* allele (see Section 5.4), and the fact that its Northern phenotype is similar to 121e supports the hypothesis that 121e is a splicing mutant. However it is difficult to interpret the observations made for 121e as the film is heavily over-exposed and therefore all comparisons made with other mutants (196a and 124c) must take this into account. It is interesting to compare the Northern blot phenotypes exhibited by 196a and SPJ8.31 (see Figs. 5.4B and 5.1A, respectively). The results presented for these mutants suggest that different alleles can produce variable Northern phenotypes and this may indicate a multifunctional protein in the splicing process. Indeed the *prp8.7* allele also exhibits a Northern phenotype distinct from that produced by the other *prp8* alleles (J. Brown, pers. comm.).

#### 5.3.3 Mutants in Which There Appears to be a Small Decrease in Splicing Efficiency

A comparison of the Northern phenotypes exhibited by the mutants identified in the screening procedure with those of DBY745 (all the mutants that could be placed in this group were derived from DBY745), 196a (*prp8*) and *prp25*, identified a set of mutants which appear to show a relatively small decrease in splicing efficiency (see Fig. 5.5). This analysis is based entirely on the Northern phenotype produced by mutants when probed with the *CYH2* gene, and indeed possible splicing defects are undetected by analysing the actin transcript (see Fig. 5.5). It can be seen that the mutants may produce a relative accumulation of pre-mRNA, as

**FIGURE 5.4A: NORTHERN BLOT HYBRIDISATION ANALYSIS OF MUTANT 121e**

- Tracks 1-2: RNA was prepared from cultures of 121e grown in YPDA at 23°C or shifted to 36°C for 1.5-2 hours. All other experimental details were as for Fig. 5.1.
- Tracks 3-6: Time-course of Northern blot phenotype exhibited by 121e on incubation at 36°C. Cultures of 121e were initially grown in YPDA at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 and then shifted to 36°C. Thereafter, RNA was prepared at hourly intervals (1-4 hours, tracks 3-6 respectively).

**FIGURE 5.4B: NORTHERN BLOT HYBRIDISATION ANALYSIS OF MUTANT 196a**

- Tracks 1-4: RNA was prepared from cultures of 196a grown in YPDA at 23°C or shifted to 36°C for 1.5-2 hours. Tracks 1 and 2 and tracks 3 and 4 represent two independent experiments. All other experimental details were as for Fig. 5.1.

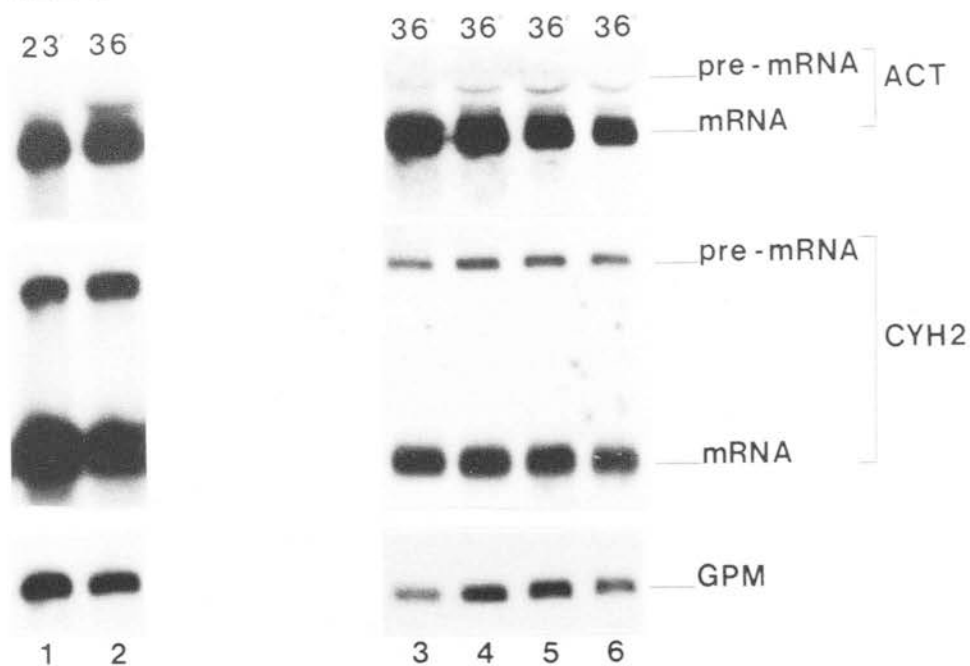
**FIGURE 5.4C: NORTHERN BLOT HYBRIDISATION ANALYSIS OF MUTANT 109i**

- Tracks 1-4: Time-course of the Northern blot phenotype exhibited by 109i on incubation at 36°C. All experimental details as for Fig. 5.4A, tracks 3-6.

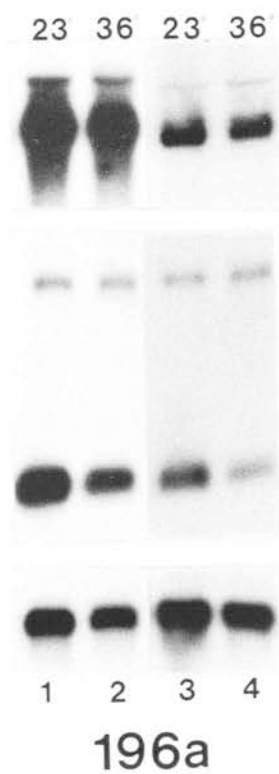
**FIG. 5.4**

**A**

**121e**



**B**



**C**

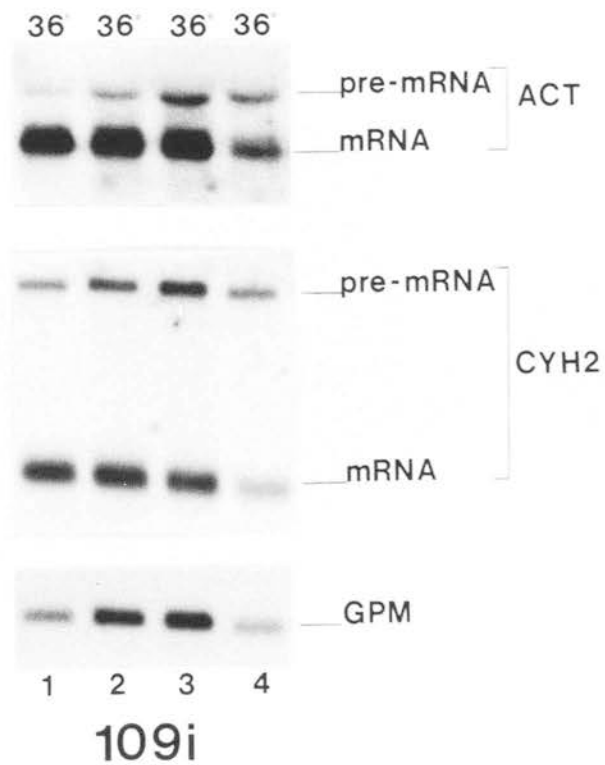
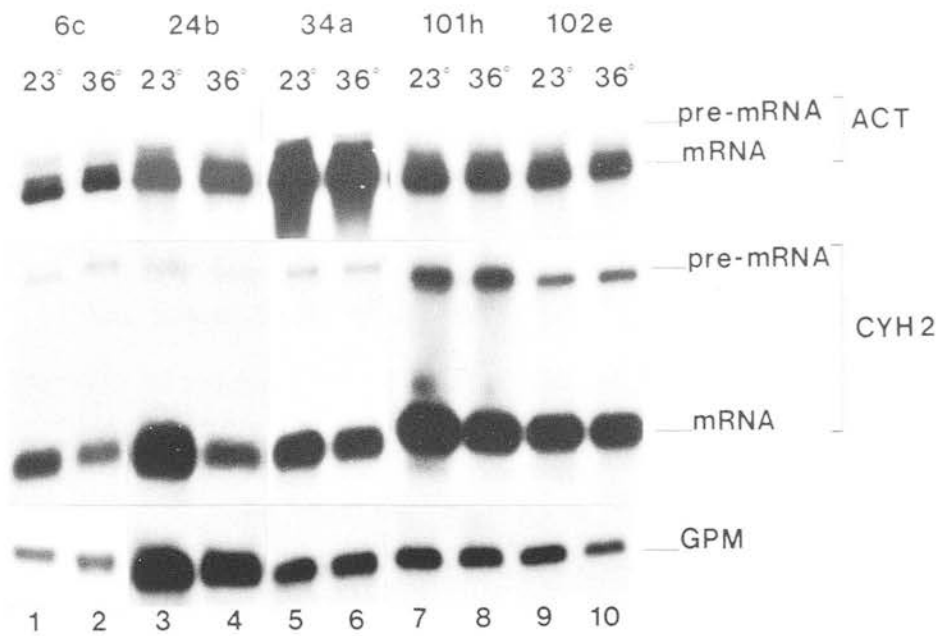


FIGURE 5.5: NORTHERN BLOT HYBRIDISATION ANALYSIS OF OTHER  
POSSIBLE *prp* MUTANTS

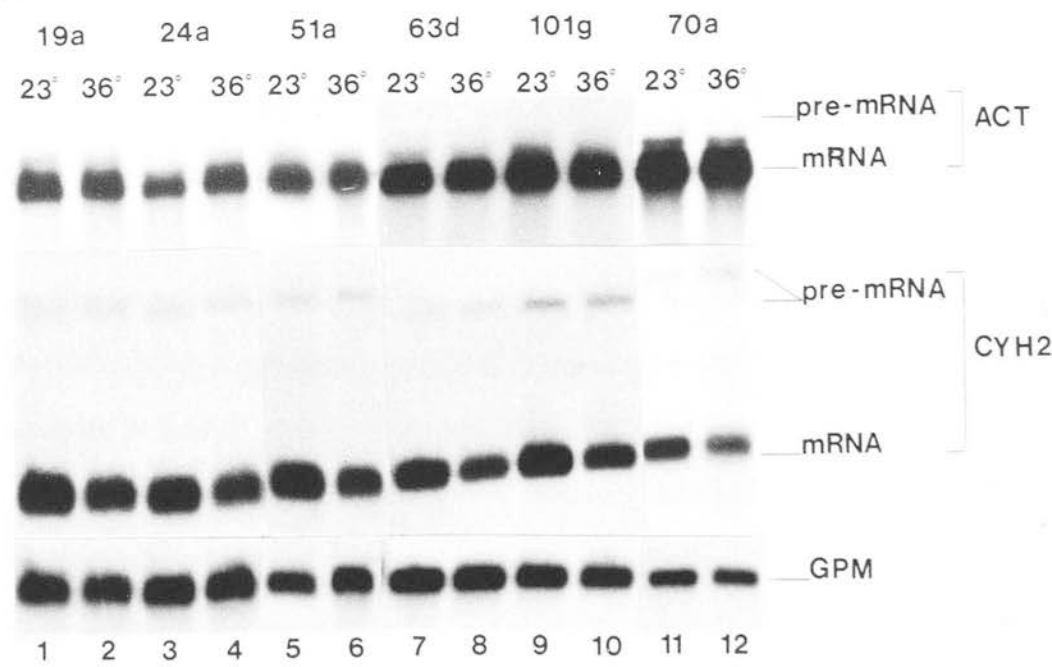
All experimental details were as for Fig. 5.1. The mutants presented in this figure represent those which may exhibit a relatively small decrease in the efficiency of splicing. The names of the mutants are indicated above the tracks.

FIG. 5.5

A



B



observed for 196a and *prp25* at 36°C (see Fig. 5.4B and 5.7A respectively). Of those indicated in Figs. 5.5A and B, the mutants 6c, 24b, 34a, 51a, 63d, 70a, 101g and 101h appear to exhibit the most pronounced contrast in pre-mRNA and mRNA ratios. Further experiments analysing different transcripts are required to determine whether these mutants are defective in splicing. Because of the ambiguity in defining these mutants as *prp* strains they are not considered in the comparative analysis with other screens discussed in Section 5.6.

#### 5.3.4 Mutants With Altered Transcript Size

Three mutants, 7c, 31a and 43c fell into this class. When incubated at 36°C, 7c exhibits an increase in the size of pre-mRNA (see Fig. 5.6A; indeed with shorter exposures the mRNA appears to exhibit a similar increase in size at 36°C, results not presented). 31a and 43c appear to show an increase in the size of the pre-mRNA at 23°C relative to 36°C, though the GPM control probably indicates that this effect is independent of the intron (see Fig. 5.6A). Primer extension analysis of the *CYH2* transcript using the JMCYH2A oligodeoxynucleotide (see Chapter 2.1.6, Table 2.4) and RNA preparations made from 7c and 31a indicate that the transcription initiation sites and the splice sites are identical in these mutants incubated at the two temperatures (see Fig. 5.6B and note that *CYH2* has at least three sites of transcription initiation), and this implies that the reason for the increased transcript size is a defect in determining the 3' end of the transcript. This could be at the level of transcriptional termination or polyadenylation. 43c was not analysed. It has previously been noted that on incubation at 36°C, a *prp1* mutant accumulates intron-less *GAL1*, 7 and 10 transcripts which are extended at their 3' ends (St. John and Davis, 1981) and this may be related to the *GAL* promoters as other genes do not exhibit such a phenomenon in *prp1* mutants (Struhl and Davis, 1981). In complementation analysis it is interesting to note that 7c could complement a *prp1* strain and thus was not defective in *PRP1* (see below). However a similar analysis has yet to be carried out with 31a and 43c. 31a appears to accumulate pre-mRNA at the expense of



## FIGURE 5.6A: NORTHERN BLOT HYBRIDISATION ANALYSIS OF *t<sup>s</sup>* MUTANTS 7c, 43c AND 31a

All experimental details were as for Fig. 5.1.

## FIGURE 5.6B: PRIMER EXTENSION ANALYSIS OF MUTANTS 7c AND 31a

RNA was prepared from cultures of 7c and 31a which had been grown in YPDA at 23°C or shifted to 36°C for 1.5-2 hours. The oligodeoxynucleotide JMCYH2A, which is complementary to a region of exon two (see Chapter 2.1.6, Table 2.4), was hybridised to 30-100µg total RNA from each preparation at 53°C for 15-30 minutes, and the primer extension reaction was then carried out as described in Chapter 2.2.8.5.

The markers are derived from a *Msp*I digest of pBR322, the fragments of which were radioactively labelled by rendering the ends blunt with radio-labelled nucleotides.

Tracks 1-2 and 5-6: 10µl of each RNA preparation (~30-50µg) was used in the primer extension reactions. The temperature at which the yeast cultures had been grown is indicated above each track.

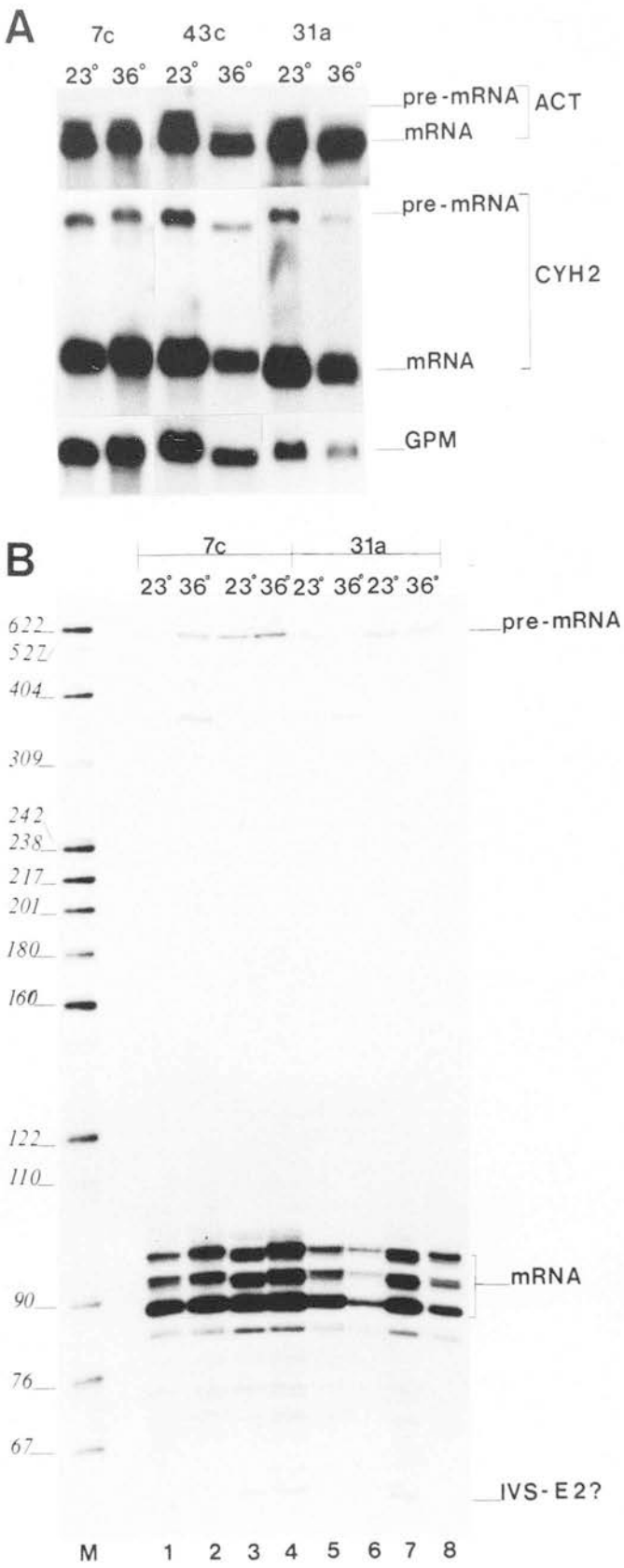
Tracks 3-4 and 7-8: 20µl of each RNA preparation (~60-100µg) was used in the primer extension reactions.

The *CYH2* gene has at least three transcription start sites, therefore the sizes of the primer extension products expected are,

pre-mRNA	-	616, 612 and 606nts (N.B. these three species have not been resolved at the top of the gel)
mRNA	-	106, 102 and 96nts
IVS-E2	-	66nts.

However the actual sizes of the pre-mRNA and mRNA species detected are roughly 10 nucleotides smaller. Using the oligodeoxynucleotide JMCYH2B, which is complementary to a region of the intron, the sizes of pre-mRNA species were also found to be roughly 10 nucleotides smaller than expected (results not presented). The reason for the discrepancy is unknown but may be due to differences in the transcription initiation sites of *CYH2* in different yeast strains.

FIG. 5.6



mRNA when incubated at 36°C and thus may be a *prp* mutant (see Fig. 5.6B, tracks 7 and 8).

The remaining mutants that originally exhibited a blue phenotype in the screening procedure fail to show a discernable *prp* phenotype on Northern blot analysis, and are not considered further.

#### 5.4 COMPLEMENTATION ANALYSIS OF POTENTIAL *prp* MUTANTS

Complementation analysis (see Chapter 2.2.3.5) was performed using some of the possible *prp* mutants isolated by the screening procedure and the results of this analysis are given in Table 5.3. All *prp* strains (indicated in Chapter 2, Table 2.2) of both mating-types were utilised in this analysis; each individual mutant was crossed at least twice to the *prp* strains to verify the results obtained. Unlinked mutations which fail to complement each other have been identified in yeast (Maddock, J., pers. comm.) and therefore further genetic analyses are required to confirm the *prp* genotype of the newly isolated mutants.

TABLE 5.3 COMPLEMENTATION ANALYSIS OF POSSIBLE *prp* MUTANTS

<u>Mutant</u>	<u>Complementation Group</u>	<u>Mutant</u>	<u>Complementation Group</u>
196a	<i>prp8</i>	60h*	-
221c	<i>prp6</i>	61d	<i>prp1</i>
275	<i>prp5</i>	104b	<i>prp24</i>
298	<i>prp6</i>	105c	<i>prp2</i>
6c*	-	109i	-
7c*	-	121e	-
9c	<i>prp24</i>	124c	<i>prp24</i>
36a	-	20a	-

(-) indicates that the mutant belongs to a new complementation group; \* indicates that Northern analysis has not been performed using diploid strains of this mutant crossed to *prp25* and *prp26*).

#### 5.4.1 Northern Analysis of Strains Crossed to *prp25* and *prp26*

Because *prp25* and *prp26* are not temperature-sensitive (see Chapter 1.9), determining whether the newly isolated mutants complement the *prp25* or *prp26* defects requires the

generation of diploid strains from the relevant matings. RNA was isolated from these diploids after their incubation at 23°C and 36°C and analysed for the defects specific to the strains by Northern blotting. Fig. 5.7A shows the Northern phenotype exhibited by *prp25* and *prp26* and the results of crosses involving 36a, 109i and 121e with *prp26*. The accumulation of the intron demonstrated by *prp26* is clearly absent from these diploids (the result for the actin intron is not presented for the cross involving 121e and *prp26*, however this diploid does not accumulate the *CYH2* intron, see tracks 9 and 10). Therefore these new strains are unlikely to be *prp26* mutants. However it is noticeable that the diploid of 121e and *prp26* appears to accumulate *CYH2* pre-mRNA (but not actin) at 36°C even though it does not accumulate the intron. A possible explanation for this could be that the relevant gene products of each strain interact to partially complement the 121e-specific defect in pre-mRNA accumulation but complement fully the defect in intron metabolism. Indeed intragenic complementation is also a possibility.

The accumulation of pre-mRNA observed for 36a, 109i and 121e is absent from crosses involving *prp25*, and it is likely that they represent a different complementation group(s) (see Fig. 5.7B). However interpretation of the crosses involving *prp25* is difficult as the individual defects are not so distinctive. Indeed for 36a and 109i crossed to *prp25* there is a definite decrease in the total RNA level at 36°C though no accumulation of pre-mRNA is detected. In this respect it is important to note that the pre-mRNA accumulation exhibited by *prp25* is not substantial (see Fig. 5.7, tracks 1 and 2) and therefore it cannot be certain that 36a and 109i are in a different complementation group to *prp25*.

It is therefore only possible to suggest that these three mutants represent at least one newly identified complementation group(s). So far the mutants have not been outcrossed and so it is not possible to determine the number of complementation groups they define. Analysis of 6c, 20a and 60h has not been completed. They are probably not *prp25* or *prp26* as they exhibit different Northern phenotypes, though the possibility exists that different mutations in the same gene exhibit different phenotypes. Indeed this is suggested by the isolation of 196a (*prp8*) and with current analysis of the *prp8.7* allele, which do not display the "normal" *prp8*

**FIGURE 5.7: NORTHERN BLOT HYBRIDISATION ANALYSIS OF CROSSES INVOLVING 36a, 109i AND 121e WITH *prp25* AND *prp26***

RNA was prepared from the relevant diploid strains grown at 23°C or shifted to 36°C for 1.5-2 hours. All experimental details were as for Fig. 5.1, except that diploid strains of 36a and 109i crossed with *prp25* and *prp26* were grown in YMM (see Chapter 2.1.2.2) plus the required supplements. In contrast diploid strains of 121e crossed with *prp25* and *prp26* were grown in YPDA. Thus the decrease in RNA levels at 36°C evident for the former diploid strains, may be a consequence of the growth medium used.

**FIGURE 5.7A**

Tracks 1-4: Northern blot analysis of *prp25* and *prp26*. The accumulation of the intron at both temperatures is clearly evident in *prp26*. The sizes of the intron species are;

actin - IVS - 308nts

*CYH2* - IVS - 510nts

Tracks 5-6 : *prp26* crossed with 36a

Tracks 7-8 : *prp26* crossed with 109i

Tracks 9-10: *prp26* crossed with 121e

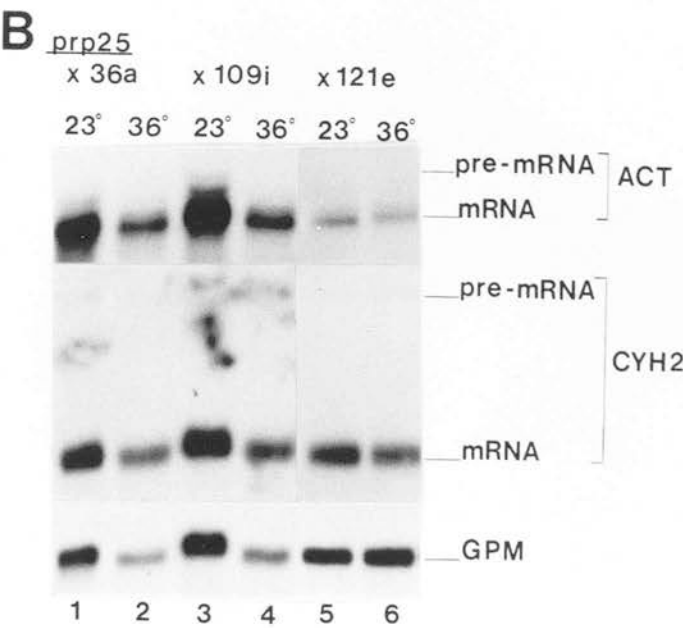
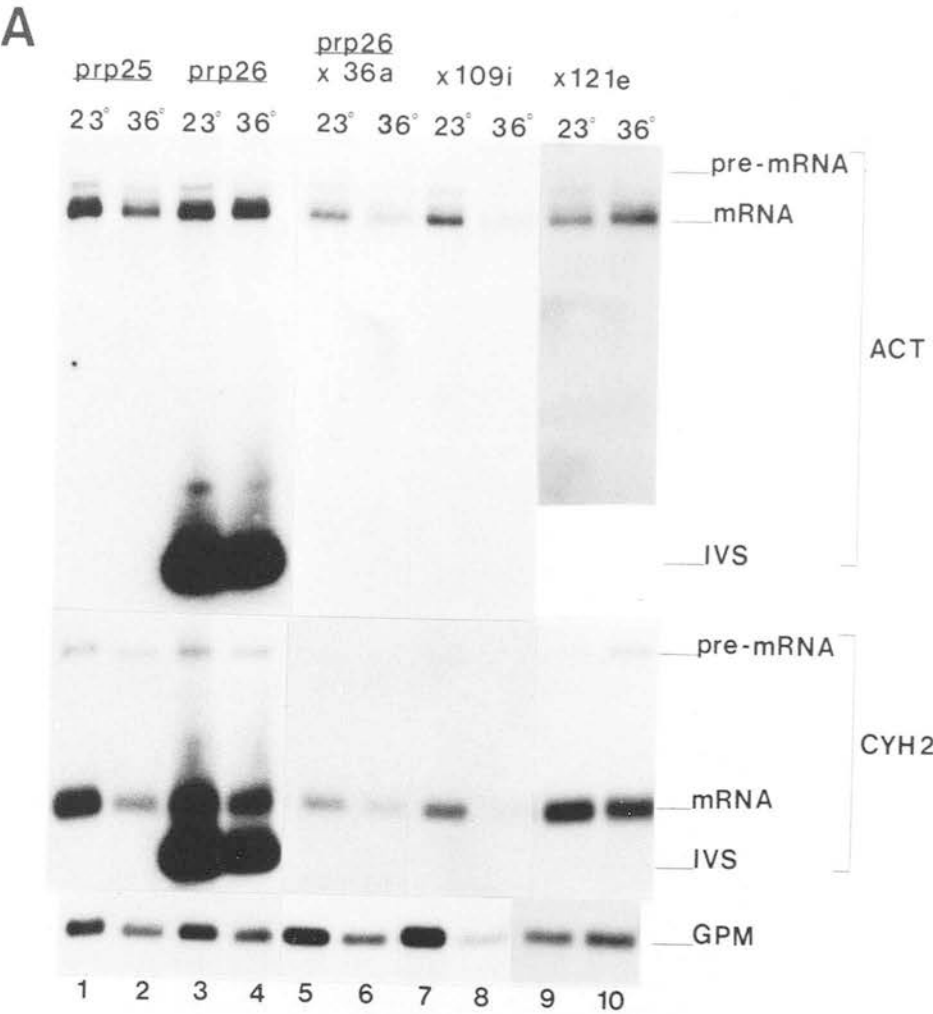
**FIGURE 5.7B**

Tracks 1-2 : *prp25* crossed with 36a

Tracks 3-4 : *prp25* crossed with 109i

Tracks 5-6 : *prp25* crossed with 121e

FIG. 5.7

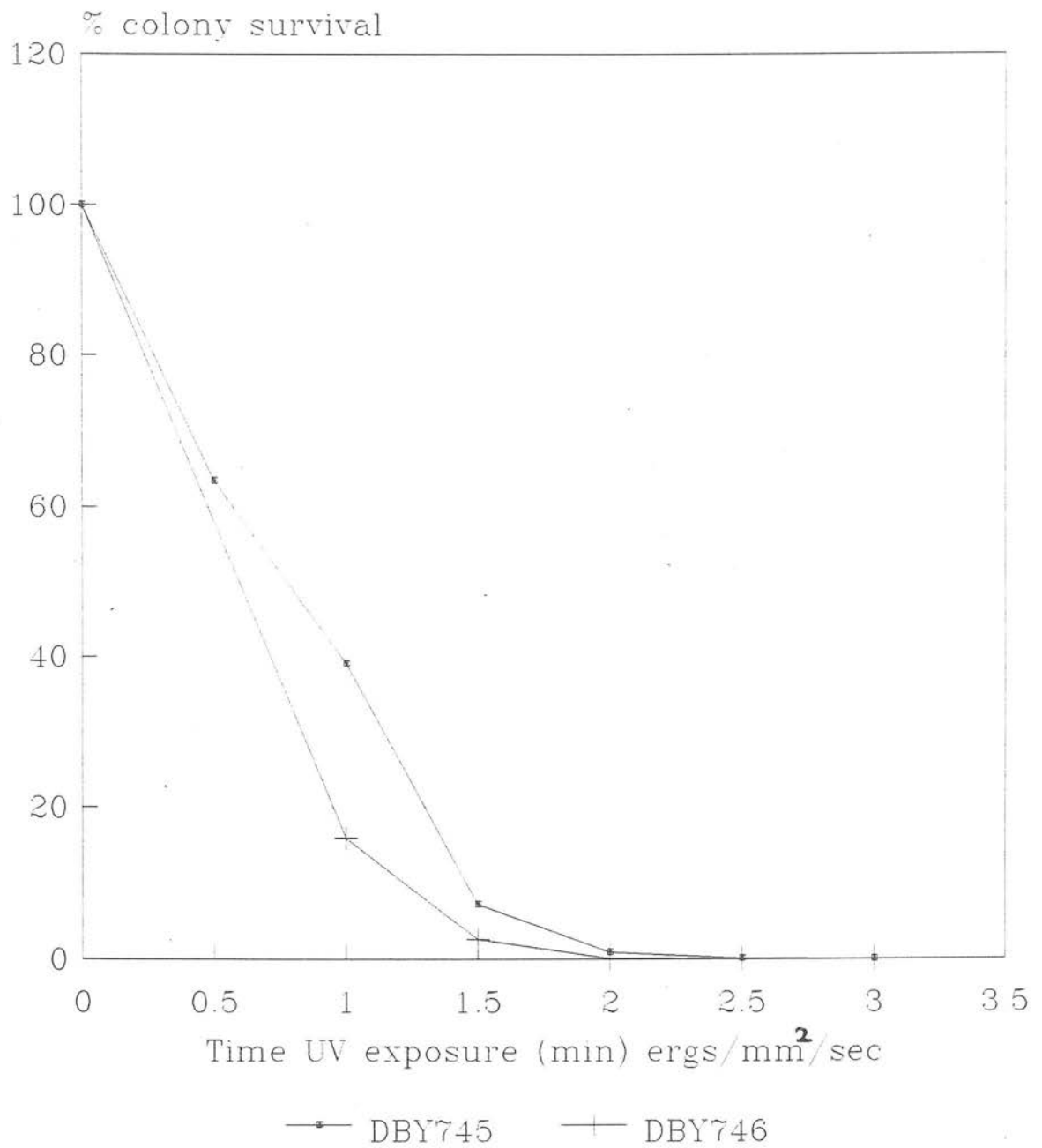


## FIGURE 5.8: UV MUTAGENESIS; SURVIVAL CURVES OF DBY756 AND DBY746

All mutagenesis procedures are described in Chapter 2.2.3.7 and 2.2.3.8. Duplicate samples of cells (for each time-point) were removed during exposure to UV light and spread on plates. The number of colonies arising from the UV treated cell-pools, after 3-5 days incubation at 23°C, was averaged and then compared to the number of colonies present from untreated samples (duplicate samples also). This comparison is expressed as the percentage survival of cells after a specific UV exposure time. The UV dose was 10 ergs/mm<sup>2</sup>/sec, delivered from a height of 32cm.



Figure 5.8  
UV Mutagenesis: Survival Curves



DBY745 and DBY746

Northern phenotype (Brown, J., pers. comm.). 124c may also be such a mutant, depending on the "normal" *prp24* Northern blot phenotype. 6c, 20a and 60h are not allelic to previously isolated *prp* mutants (depending on the analysis with *prp25* and *prp26*). 7c was analysed to determine if it could complement a *prp20* mutant (Vijayraghavan *et al.*, 1989) or a *prp1* mutant (see above), but it proved to be a new complementation group (depending on analysis with *prp25* and 26).

## 5.5 *t<sup>s</sup>* PROFILE OF ISOLATED MUTANTS

All mutants which exhibited a blue phenotype, as assayed in the screening procedure, (see Chapter 4, Fig. 4:10) were resuspended in water and spotted onto YMG/Cas and YPDA plates and incubated at 23°C for 4 days. The cells were then replica-plated onto the same media used for its initial growth at 23°C and incubated at various temperatures for 5 days. The results of the growth phenotypes at each temperature are given in Table 5.4. Only those mutants which constitute the 4 groups in Section 5.3 are indicated.

TABLE 5.4 *t<sup>s</sup>* GROWTH PROFILE OF POSSIBLE *prp* MUTANTS

Mutant	18°C		23°C		30°C		34°C		36°C	
	YMG	YPDA	YMG	YPDA	YMG	YPDA	YMG	YPDA	YMG	YPDA
9c	+	+	+	+	+/-	+/-	-	-	-	-
20a	+	-	+	+	+	+	+	+	-	-
36a	+	+	+	+	+	+	+	+/-	-	-
60h	+	+	+	+	+	+	+	+/-	-/+	-
61d	+	+	+	+	+	+	+	-	-	-
104b	+	+	+	+	+	+	+	+	-	-
105c	+	+	+	+	+	+	+	+	-	-
109i	+	+	+	+	+	+	+	+	-	-
121e	+	+	+	+	+	+	+	-	-	-
124c	+	+	+	+	+	+	+	+	-	-
196a	+	+	+	+	+	+	+/-	+/-	-	-
221c	+	+	+	+	-	-	-	-	-	-
275	+	+	+	+	+	+	-	-	-	-
298	+	+	+	+	+/-	+/-	-	-	-	-
6c	+	+	+	+	+	+	+	+	-	-
19a	+	+	+	+	+	+/-	-	-	-	-
24a	+	+	+	+	+	+	+	+	-/+	-/+
24b	+	+	+	+	+	+	-	-	-	-
34a	+	+	+	+	+	+	+	+	-/+	-/+
51a	+	+	+	+	+	+	+	+	-	-
70a	+	+	+	+	+	+	+/-	+/-	-	-
101h	+	+	+	+	+	+	+	+	-	-
102e	+/-	+/-	+/-	+/-	+	+/-	+	+	+/-	-
110a	+	+	+	+	+	+	+	+	-	-
7c	+	+	+	+	+	+	+	+	-	-
31a	+	+	+	+	+	+	+	+	-	-
43c	+	+	+	+	+	+/-	+/-	-	-	-

(+) indicates growth; (-) indicates no growth; (+/-) indicates partial growth which could be a result of a partial *t<sup>s</sup>* defect, reversion and/or contamination of the original water-suspended colony)

Results for the all the  $r^s$  mutants isolated are not presented. These growth profiles are tentative and need to be repeated. At the time of writing a thorough investigation of the genotypes of all the mutants has not been performed, and it is therefore feasible that some may be yeast contaminants e.g. 102e. The reason(s) for the discrepancy in growth occasionally observed between YMG/Cas and YPDA is unknown.

## 5.6 DISCUSSION

### 5.6.1 Comparison of the Efficiency of the Plate-Screening Procedure Using DBY745 and DBY746

Out of 296  $r^s$  mutants derived from DBY745, five were new isolates of pre-existing *prp* mutants (9c, 61d, 104b, 105c and 124c) and five exhibited properties that suggested that they were newly defined *prp* complementation groups (20a, 36a, 60h, 109i and 121e). Thus the *prp* mutants accounted for 3.38% of the total  $r^s$  pool (see Section 5.6.2 for a comparison with other screens). However it has not been shown that 20a and 36a can consistently exhibit a Northern blot phenotype that would indicate that they were *prp* mutants. Thirty-five  $r^s$  mutants produced a blue phenotype in the plate-screening procedure and ten of these mutants exhibited a possible *prp* phenotype when investigated by Northern blot analysis. Thus 28.6% of the mutants screened by Northern blot analysis were possible *prp* mutants.

Out of 294  $r^s$  mutants derived from DBY746, four were allelic mutants of pre-existing *prp* complementation groups (196a, 221c, 275 and 298). Thus the *prp* mutants accounted for 1.36% of the total  $r^s$  pool (it is possible that 196b is also in the same complementation group as *prp8* because it was derived from the same mutagenised cell-pool as 196a; however complementation analysis has not been performed using this mutant). Eleven mutants produced a blue phenotype in the plate-screening method and on Northern blot analysis 36.36% of these were *prp* mutants.

It would appear from these observations that it is easier to specifically identify *prp* mutants when DBY746 is the strain utilised in the plate-screening procedure (a 36.36% success rate compared to one of 28.6%, at best, for DBY745). A possible explanation for this difference is the fact that DBY745 produces a considerable amount of background  $\beta$ -galactosidase activity (see Chapter 4.1.2). Using the crude plate assay this makes it difficult to distinguish the relatively small increase in activity produced by the weak *prp* mutants from the background activity exhibited by most of the other colonies, and this results in the "positive" blue phenotype

of many more colonies. It should be noted that colonies that were slightly blue were isolated and screened by Northern blot analysis. However it is also clear that mutants that show a weak *prp* phenotype, as investigated by Northern blot analysis (e.g. 61d, 60h and 109i), are clearly under-represented in the screen utilising DBY746. Indeed if 60h, 61d, 109i, 20a and 36a are excluded from the final tally of possible *prp* mutants isolated using DBY745, then the percentage of *prp* mutants isolated using both strains is very similar (1.36% for DBY746 to 1.69% for DBY745). DBY746 is *t<sup>s</sup>* for galactose induction and it is possible that this restricts the types of *prp* mutants that can be isolated using this strain (see Chapter 4.2.1.3 for a fuller explanation). It seems likely that the screen would function with increased efficiency if a suitable strain was utilised that was not *t<sup>s</sup>* for galactose induction and which did not produce high background levels of activity.

### 5.6.2 Comparison of the Plate-Screening Method With Other Screening Procedures

Two other large-scale screening procedures have identified conditional-lethal *prp* mutants (see Chapter 1.9). The first screen defined mutants defective in RNA metabolism on the basis that they exhibited a protein:RNA ratio  $>4$  ( $P/R >4$ ; henceforth called the P/R screen) when incubated at the non-permissive temperature (Hartwell, 1967) and the second identified *prp* mutants by Northern blot analysis as described in Section 5.3 and in Chapter 1.9 (Vijayraghavan *et al.*, 1989). *prp* mutants isolated by each screen accounted for roughly 5% and 3.7% respectively of the *t<sup>s</sup>* pools generated.

#### 5.6.2.1 Advantages of the Plate-Screening Procedure

The plate-screening method is significantly faster than either the P/R or Northern blot screens. It reduces the number of *t<sup>s</sup>* mutants that have to be analysed by Northern blotting procedures almost 13-fold (the average of DBY745 and DBY746) and indeed this may be further reduced depending on the strain utilised in the screen (see 5.6.1), e.g. if DBY746 is considered on its own then there is a 30-fold reduction in the number of mutants that were analysed using Northern blotting techniques. In practice the rate-limiting step in the whole procedure (from the generation of *t<sup>s</sup>* mutants to the Northern blot analysis of potential *prp* mutants) is the creation of a *t<sup>s</sup>* mutant pool.

The sensitivity of the plate-assay can be modulated by simply changing the

concentration of X-gal in the reaction (see Chapter 4.2.3). Thus the mutants can be induced on YMR/Gal/Cas medium for relatively short or long incubations at the appropriate temperatures. In this way it should be possible to differentiate those mutants which take longer to exhibit an effect on splicing e.g. snRNP assembly mutants, from the normal wild-type background activity. The approaches taken using the other two screening procedures may not have identified such mutants as the  $t^s$  mutants were analysed after only 2-3 hours incubation at the non-permissive temperature (Hartwell, 1967; Vijayraghavan *et al.*, 1989). SnRNP particles are metabolically stable and therefore it is probable that mutations in snRNP assembly would take longer to produce an effect on splicing. It would not be unreasonable to expect that such mutants would take longer to exhibit any  $t^s$  effect on their rate of growth. Indeed 196a only starts to show a temperature effect on its rate of growth (as measured by the OD 600<sub>nm</sub>) after 10 hours incubation at 36°C (results not presented). Experiments are underway to examine the  $t^s$  growth profile of all the mutants isolated by this screen. In this respect it is interesting to note that 20a and 36a have yet to exhibit a reproducible accumulation of pre-mRNA, and it may be that they require extended lengths of incubation at 36°C i.e. >4 hours; indeed 36a accumulates pre-mRNA after 4 hours incubation at 36°C (see 5.3.2.2).

#### 5.6.2.2 Disadvantages of the Plate-Screening Procedure

The plate-screening procedure relies on the accumulation of pre-mRNA that can encode an active  $\beta$ -galactosidase fusion protein. Consequently it may not be able to detect mutants that accumulate either the splicing intermediates and/or products only (see Vijayraghavan *et al.*, 1989). Indeed no such mutants were detected in this particular screen. However it is possible that mutations which cause accumulation of the intermediates and/or products of splicing may feedback inhibit earlier steps in the splicing reaction by titrating out snRNPs and other factors required for the initial stages in splicing. This would result in an accumulation of pre-mRNA which should be detectable using the plate-assay. This postulated titration of snRNPs and other factors may take a relatively long time to produce an effect on pre-mRNA accumulation, and therefore it is of benefit that the plate-screening procedure can be prolonged by regulating the sensitivity of the assay. It is therefore possible that such mutants were not detected using this plate-screening assay because a relatively small pool of  $t^s$  mutants was analysed. However it must be noted that in the Northern blot screen of Vijayraghavan *et al.*, (1989), nine out of the thirty-seven *prp* mutants isolated accumulated the intermediates or products of splicing. It would be expected that a similar proportion would have been detected using the plate-assay screen. These differences could be a result of the screen employed (i.e this

screen relies on the accumulation of pre-mRNA, see above) or a consequence of the different forms of mutagenesis utilised (i.e. UV mutagenesis versus chemical mutagenesis).

The number of *prp* mutants isolated by the plate-screening procedure, expressed as a percentage of the number of mutants in the  $r^s$  pool, may be lower than those identified by the other two methods e.g. 3.38% when DBY745 is used in the plate- screen compared to ~5% and 3.7% for the P/R and Northern screens respectively. It is interesting to note that in the Northern screen if the mutants that accumulate intermediates or products only are excluded from this analysis then the percentage recovery of *prp* mutants falls to ~2.8%. Thus the differences in the percentage recovery of *prp* mutants may be due to the possibility that the plate-screen cannot detect mutants which accumulate the intermediates or products of splicing (see above). However the figures for each screening procedure are relatively close. It should be noted that a minority of the mutants originally placed in the  $r^s$  pool created in this study were not infact temperature- sensitive and therefore this may reflect the slight decrease in the percentage recovery of mutants observed for the plate-screen when DBY745 was used. The reason for the relatively low recovery of *prp* mutants using DBY746 may be a consequence of its temperature-sensitive galactose induction phenotype (see above).

#### 5.6.3 New Isolates of Pre-Existing *prp* Mutants Isolated by the Plate-Assay Screening Procedure

Using the plate-assay, alleles of *prp1* (one isolate), *prp2* (one isolate), *prp5* (one isolate), *prp6* (two isolates), *prp8* (one isolate) and *prp24* (three isolates) were identified. It has been proposed that mutations at the commitment stage in splicing may allow the escape of relatively more pre-mRNA from the nucleus (Legrain and Rosbash, 1989). The plate-assay screen would therefore readily identify "commitment" mutants by their increased  $\beta$ -galactosidase activities. PRP6 may be involved in the commitment stage of splicing (see Legrain and Rosbash, 1989) and it is interesting to note that two independent *prp6* alleles were isolated using the plate-assay screen. It has also been suggested that PRP24 is involved in the early stages of spliceosome assembly (Arenas *et al.*, 1989) and indeed three independent *prp24* alleles were identified in this screen. The plate-assay screen may therefore enhance the detection of mutants defective in the early stages of splicing.

The new *prp8* allele (196a) exhibits a Northern blot phenotype unlike the other *prp8* alleles (except *prp8.7* which does not readily accumulate pre-mRNA unless extended incubation, >8 hours at 36°C, is undertaken). Consistent with the observation that the mRNA level is not

significantly depleted on incubation at 36°C for 2 hours, 196a continues to grow for up to 10 hours at the non-permissive temperature (see Section 5.6.2.1). The fact that the growth of 196a is temperature-sensitive suggests that the effect of the mutation is more pronounced at the higher temperature. After only two hours incubation 196a accumulates pre-mRNA at both 23°C and 36°C. The early accumulation of pre-mRNA may not be consistent with 196a being defective in snRNP assembly. However it could be that snRNP assembly is affected at both temperatures producing partially active snRNPs which can function in splicing. In this case the defect would be greater at the higher temperature. These observations may suggest that 196a is defective in the splicing process directly, rather than in snRNP assembly.

Isolation of a *prp1* allele which accumulated pre-mRNA was unexpected. Based on the results presented in Chapter 4.1.2.4 it was expected that *prp1* mutants would not be isolated using the plate-assay screen. PRP1 has been implicated in a variety of RNA processing events. Early experiments suggested that *prp1* mutants were defective in the transport of RNA from the nucleus (Hutchison *et al.*, 1969; Shiokawa and Pogo, 1974). It was subsequently observed that the *prp1-1* mutant was also defective in the processing of some pre-tRNAs and pre-rRNAs when incubated at the restrictive temperature (Hopper *et al.*, 1978; Knapp *et al.*, 1978). It was also found that the levels of intron-containing and intron-less transcripts were reduced in a similar manner when a *prp1-1* mutant was incubated at 36°C (Rosbash *et al.*, 1981). This suggested that PRP1 was not directly involved in pre-mRNA processing. The *PRP1* gene has been cloned and sequenced and deletion analysis has indicated that a specific region of the *PRP1* protein is required to produce an effect on pre-tRNA or pre-rRNA processing (Atkinson *et al.*, 1985; Traglia *et al.*, 1989). The new *prp* allele isolated in the plate-assay screen constitutes the first evidence implicating PRP1 in pre-mRNA processing. However it is possible that this *prp1* mutant also contains another *prp* mutation which is responsible for the accumulation of pre-mRNA observed in this mutant. If the introduction of the *PRP1* gene (on a yeast expression vector) into this mutant restores the ability of the cells to splice pre-mRNA then this will indicate that the accumulation of pre-mRNA is due to the *prp1* mutation. This experiment is currently underway.

#### 5.6.4 Other Types of Mutants that Could be Isolated Using the Plate-Assay Screen

Mutations that caused an aberrant 5' or 3' splice site cleavage event could lead to the splicing of the pJBM-4 fusion- transcript back into frame with the *lacZ* gene. This mRNA



transcript would then be able to encode an active  $\beta$ -galactosidase fusion protein. The plate-assay would, in theory, detect this novel class of splicing mutants.

Mutants which exhibited altered rates of translation or could change translational reading frames could possibly be isolated using this screen.

Mutants which were defective in RNase activity would produce relatively more pre-mRNA than those which were not, thus becoming detectable using this screen.

#### 5.6.5 Future Work

The isolation of potential *prp* mutants which may represent new complementation groups was achieved using a novel screening procedure. It is now important to design experiments that can determine whether these mutants are defective in the process of pre-mRNA splicing. Yeast provides a powerful system in which various complementary approaches can be taken to study the function of gene products.

Initially these mutants will be out-crossed into the appropriate wild-type genetic background. Tetrad analysis will be performed to establish a 2:2 segregation pattern of the *i<sup>s</sup>* mutation. It will also be ascertained whether the RNA splicing defect co-segregates with the *i<sup>s</sup>* phenotype. Following this the mutants will be crossed with each other to identify the number of new complementation groups isolated. An intensive study of the *i<sup>s</sup>* growth profile and pre-mRNA accumulation pattern of these mutants will then be initiated (see Section 5.6.2.1). Complementation of the *i<sup>s</sup>* defect in these mutants should enable the isolation of the wild-type gene of interest. The cloning of the wild-type genes should allow the generation of antibodies directed against their protein products (if indeed they are protein-encoding genes) and thus facilitate a number of experiments which could determine the role of these proteins in splicing (e.g. see Lossky *et al.*, 1987). Other approaches designed to determine the role of the *prp* products in splicing include making heat-sensitive splicing extracts (Lustig *et al.*, 1986) and mutational analysis of the genes themselves.

#### 5.6.6 Other Possible Screens

Using yeast strains harbouring the plasmid pJBM-4/s it should be possible to use the same plate-assay procedure as described in Chapter 4, Fig. 4.10, to detect conditional-lethal mutants that are defective in the choice of the correct 5' and 3' splice sites when incubated at the restrictive temperature. Using an appropriate strain the background activity arising from pJBM-4/s could be considerably reduced (see Chapter 4.1.2.1). *Trans*- or *cis*-acting mutations

could conceivably cause an alternative splicing event generating mRNA transcripts (derived from pJBM-4/s) that are able to encode an active  $\beta$ -galactosidase fusion protein. Because the turnover rate of a mRNA transcript is considerably less than that of the pre-mRNA and because in general pre-mRNA may be retained in the nucleus whereas mRNA is transported out, it is likely that the detection of these splice site mutants would be enhanced in a screen where it is the mRNA that encodes the  $\beta$ -galactosidase activity rather than the pre-mRNA, as for pJBM-4. Thus the detection of conditional-lethal 5' and 3' splice site mutants may be relatively straightforward.

## REFERENCES

- Abelson, J., (1979). *Ann. Rev. Biochem.*, 48: 1035-1069.
- Aebi, M., Hornig, H., Padgett, R.A., Reiser, J., and Weissmann, C. (1986). *Cell*, 47: 555-565.
- Aebi, M., Hornig, H., and Weissmann, C. (1987). *Cell*, 50: 237-246.
- Agris, C.H., Nemeroff, M.E. and Krug, R.M. (1989). *Mol. Cell. Biol.*, 9, 259-267.
- Alibert, C., Tozi, J., Temsamani, J., Jeanteur, P., Brunel, C. and Cathala, G. (1990). *Nucleic Acids Res.*, 18: 235-245.
- Anderson, G.J., Bach, M., Luhrmann, R. and Beggs, J.D. (1989). *Nature* 342: 819-821.
- Arenas, J., Company, M., Vijayraghavan, U. and Abelson, J. (1989). *RNA Processing Meeting Abstract Book*, Cold Spring Harbor, pp. 31.
- Ares, M, Jr. (1986). *Cell*, 47: 49-59.
- Aronow, B., Lattier, D., Silbiger, R., Dusing, M., Hutton, J., Jones, G., Stock, J., McNeish, J., Potter, S., Witte, D. and Wiginton, D. (1989). *Genes Dev.*, 3: 1384-1400.
- Atkinson, N.S., Dunst, R.W. and Hopper, A.K. (1985). *Mol. Cell. Biol.*, 5: 907- 915.
- Augustin, S., Müller, M.W. and Schweyen, R.J. (1990). *Nature* 343: 383-386.
- Bach, M., Winkelmann, G. and Luhrmann, R., (1989). *Proc. Natl. Acad. Sci. USA*, 86: 6038-6042.
- Bach, M., Krol, A. and Luhrmann, R. (1990). *Nucleic Acids Res.*, 18: 449-457.
- Baim, S.B. and Sherman, F. (1988). *Mol. Cell. Biol.*, 8: 1591-1601.
- Banroques, J. and Abelson, J.N. (1989). *Mol. Cell. Biol.*, 9: 3710-3719.
- Barabino, S.M.L., Sproat, B.S., Ryder, U., Blencowe, B.J. and Lamond, A.F. (1989). *EMBO J* 8: 4171-4178.
- Beggs, J.D., van den Berg, J., van Ooyen, A. and Weissmann, C. (1980). *Nature* 283: 835-840.
- Bell, L.R., Maine, E.M., Schedl, P. and Cline, T.W. (1988). *Cell*, 55: 1037-1046.
- Berget, S.M. and Robberson, B.L. (1986). *Cell*, 46: 1986-1996.
- Berget, S.M., Moore, C. and Sharp, P.A. (1977). *Proc. Natl. Acad. Sci. USA*, 74: 3171-3175.

- Berman, S.A., Bursztajn, S., Bowen, B. and Gilbert, W. (1990). *Science*, 247: 212-214.
- Beyer, A.L. and Osheim, Y.N. (1988). *Genes Dev.*, 2: 754-765.
- Bindereif, A., and Green, M.R. (1986). *Mol. Cell. Biol.*, 6: 2582-2592.
- Bindereif, A. and Green, M.R. (1987). *EMBO J.*, 6: 2415-2424.
- Bindereif, A., Wolff, T. and Green, M.R. (1990). *EMBO J.*, 9: 251-255.
- Birnboim, H.C. and Doly, J. (1979). *Nucleic Acids Res.*, 7: 1513-1523.
- Black, D.L., Chabot, B. and Steitz, J.A. (1985). *Cell*, 42: 737-750.
- Black, D.L., and Steitz, J.A. (1986). *Cell*, 46: 697-704.
- Black, D.L., and Pinto, A.L. (1989). *Mol. Cell. Biol.*, 9: 3350-3359.
- Blake, C.C.F. (1978). *Nature*, 273: 267.
- Blencowe, B.J., Sproat, B.S., Ryder, U., Barabirio, S. and Lamond, A.I. (1989). *Cell*, 59: 531-539.
- Boyer, H.W. and Roulland-Dussoix, D. (1969). *J. Mol. Biol.*, 41: 459-472.
- Bozzoni, I., Annesi, F., Beccari, E., Fragapane, P., Pierandrei-Amaldi, P. and Amaldi, F. *J. Mol. Biol.*, 180: 1173-1178.
- Breitbart, R.E., Andreadis, A. and Nadal-Ginard, B. (1987). *Ann. Rev. Biochem.*, 56: 467-495.
- Bringmann, P., Rinke, J., Appel, B., Reuter, R. and Lührmann, R. (1983). *J. Biol. Chem.*, 258: 2745-2747.
- Bringmann, P., Appel, B., Rinke, J. Reuter, R., Theissen, H. and Lührmann, R. (1984). *EMBO J.*, 1: 1259-1265.
- Bringmann, P. and Lührmann, R. (1986). *EMBO J.*, 5: 3509-3516.
- Brody, E. and Abelson, J. (1985). *Science*, 228: 963-967.
- Brow, D.A. and Guthrie, C. (1988). *Nature* 334: 213-218.
- Brow, D.A. and Guthrie, C. (1989). *Nature* 337: 14-15.
- Brown, A.J.P. (1989). In: *Molecular and Cell Biology of Yeasts* (eds) E.F. Walton and G. T. Yarranton. Blackie, pp. 70-106.

- Buchmann, A.R. and Berg, P. (1988). *Mol. Cell. Biol.*, 8: 4395-4405.
- Burgess, S., Couto, J.R. and Guthrie, C. (1990). *Cell*, 60: 705-717.
- Busch, H., Reddy, R., Rothblum, L. and Choi, Y.C. (1982). *Ann. Rev. Biochem.*, 51: 617-654.
- Callis, J., Fromm, M. and Walbot, V. (1987). *Genes Dev.*, 1: 1183-1200.
- Carlock, L. and Jones, N.C. (1981). *Nature* 294: 572-574.
- Casadaban, M.J., Chou, J. and Cohen, S.N. (1980). *J. Bact.*, 143: 971-980.
- Cech, T.R. and Bass, B.L. (1986). *Ann. Rev. Biochem.*, 55: 599-629.
- Cellini, A., Felder, E. and Rossi, J.J. (1986a). *EMBO J.*, 5: 1023-1030.
- Cellini, A., Parker, R., McMahon, J., Guthrie, C. and Rossi, J. (1986b). *Mol. Cell. Biol.*, 6: 1571-1578.
- Chabot, B. and Steitz, J.A. (1987a). *Mol. Cell. Biol.*, 7: 281-293.
- Chabot, B. and Steitz, J.A. (1987b). *Mol. Cell. Biol.*, 7: 698-707.
- Chabot, B., Black, D.L., Lemaster, D.M. and Steitz, J.A. (1985). *Science* 230: 1344-1349.
- Chang, T-H., Clark, M.W., Lustig, A.J., Cusick, M.E. and Abelson, J. (1988). *Mol. Cell. Biol.*, 8: 2379-2393.
- Chebli, K., Gattoni, R., Schmitt, P., Hildwein, G. and Sterenin, J. (1989). *Mol. Cell. Biol.*, 9: 4852-4861.
- Cheng, S-C. and Abelson, J. (1986). *Proc. Natl. Acad. Sci. USA*, 83: 2387-2391.
- Cheng, S-C. and Abelson, J. (1987). *Genes Dev.*, 1: 1014-1027.
- Choi, Y.D., Grabowski, P.J., Sharp, P.A. and Dreyfuss, G. (1986). *Science* 231: 1534-1539.
- Christofori, G., Frendewey, D. and Keller, W. (1987). *EMBO J.*, 6: 1747-1755.
- Chung, S. and Perry, R.P., (1989). *Mol. Cell. Biol.*, 9: 2075-2082.
- Ciejek, E.M., Nordstrom, J.L., Tsai, M-J. and O'Malley, B.W. (1982). *Biochem. J.*, 21: 4945-4953.
- Cigan, A.M. and Donahue, T.F. (1987). *Gene* 59: 1-18.

- Conway, G.C., Krainer, A.R., Spector, D.L. and Roberts, R.J. (1989). *Mol. Cell. Biol.*, 9: 5273-5280.
- Couto, J.R., Tamm, J., Parker, R. and Guthrie, C. (1987). *Genes Dev.*, 1: 445-455.
- Crick, F. (1978). *Science* 204: 264-271.
- Dabeva, M.D., Post-Beittenmiller, M.A. and Warner, J.R. (1986). *Proc. Natl. Acad. Sci. USA*, 83: 5854-5857.
- Dalrymple, M.A., Petersen-Björn, Friesen, J.D. and Beggs, J.D. (1989). *Cell*, 58: 811-812.
- Darnell, J.E., Jr. (1978). *Science*, 202: 1257-1260.
- de la Pena, P. and Zasloff, M. (1987). *Cell*, 50: 613-619.
- De Robertis, E.M. (1983). *Cell*, 32: 1021-1025.
- Dibb, N.J. and Newman, A.J. (1989). *EMBO J.* 8: 2015-2021.
- Doolittle, W.F. (1978). *Nature* 272: 581-582.
- Doolittle, W.F. and Sapienza, C. (1980). *Nature* 284: 601-603.
- Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E. and Abelson, J. (1984). *Cell*, 39: 611-621.
- Duchêne, M., Löw, A., Schwerzer, A. and Domdey, H. (1988). *Nucl. Acids Res.*, 16: 7233-7239.
- Dujon, B. (1989). *Gene*, 82: 91-113.
- Eperon, L.P., Estibeiro, J.P. and Eperon, I.C. (1986). *Nature* 324: 280-282.
- Fabrizio, P., McPheeters, D.S. and Abelson, J. (1989). *Genes Dev.*, 3: 2137-2150.
- Feinberg, A.P., and Vogelstein, B. (1984). *Anal. Biochem.*, 137: 266-267.
- Felber, B.K., Orkin, S.H. and Hamer, D.H. (1982). *Cell*, 29: 895-902.
- Fields, C. (1990). *Nucl. Acids Res.*, 18: 1509-1512.
- Fisher, D.E., Conner, G.E., Reeves, N.H., Wisniewolski, R. and Blobel, G. (1982). *Cell*, 42: 751-758.
- Fouser, L.A., and Friesen, J.D. (1986). *Cell*, 45: 81-93.
- Fouser, L.A., and Friesen, J.D. (1987). *Mol. Cell. Biol.*, 7: 225-230.

- Fradin, A., Jove, R., Hemenway, C., Keiser, H.D., Manley, J.L. and Prives, C. (1984). *Cell*, 37: 927-936.
- French, B.T., Schumm, D. and Webb, T. (1987). *Proc. Natl. Acad. Sci. USA*, 84: 5163-5166.
- Friendewey, D. and Keller, W. (1985). *Cell* 42: 355-367.
- Fu, J.L. and Manley, J.L. (1987). *Mol. Cell. Biol.*, 7: 738-748.
- Furueaux, H.M., Perkins, K.K., Freyer, G.A., Arenas, J. and Hurwitz, J. (1985). *Proc. Natl. Acad. Sci. USA*, 82: 4351-4355.
- Gallwitz, D. and Sures, I. (1980). *Proc. Natl. Acad. Sci. USA*, 77: 2546-2550.
- Gallwitz, D. (1982). *Proc. Natl. Acad. Sci. USA*, 79: 3493-3497.
- Gerke, V. and Steitz, J.A. (1986). *Cell* 47: 973-984.
- Ghosh, P., Roy, P., Barkan, A., Mertz, J.E. Weissman, S.M. and Lebowitz, P. (1981). *Proc. Natl. Acad. Sci. USA*, 78: 1386-1390.
- Gilbert, W. (1978). *Nature*, 271: 501.
- Gilbert, W. (1985). *Science*, 228: 823-824.
- Gilbert, W. (1986). *Nature*, 319: 618.
- Gilmartin, G.M., McDevitt, M.A. and Nevins, J.R. (1988). *Genes Dev.*, 2: 578-587.
- Goldenberg, C.J. and Raskas, H.J. (1981). *Proc. Natl. Acad. Sci. USA*, 78: 5430-5434.
- Goldenberg, C.J. and Hauser, S.D. (1983). *Nucl. Acids Res.*, 11: 1337-1348.
- Goux-Pelletan, M., Libri, D., d'Aubenton-Carafa, Y., Fozman, M., Brody, E. and Maire, J. (1990). *EMBO J.* 9: 241-249.
- Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984). *Cell*, 37: 415-427.
- Grabowski, P.J., Seiler, S.R. and Sharp, P.A. (1985). *Cell*, 42: 345-352.
- Grabowski, P.J. and Sharp, P.A. (1986). *Science* 233: 1294-1299.
- Green, M.R., Maniatis, T. and Melton, D.A. (1983). *Cell*, 32: 681-694.
- Green, M. (1986). *Ann. Rev. Genet.*, 20: 671-708.

- Gruss, P., Lai, C.-J., Dhar, R. and Khoury, G. (1979). *Proc. Natl. Acad. Sci. USA*, 76: 4317-4321.
- Guthrie, C., Riedel, N., Parker, R., Swerdlow, H. and Patterson, B. (1986a) In: *Yeast Cell Biology*. Alan R. Liss, Inc., pp. 301-321.
- Guthrie, C. (1986b). *TIBS*, 11 Oct, 430-434.
- Guthrie, C. and Patterson, B. (1988). *Ann. Rev. Genet.*, 22: 387-419.
- Hamer, D.H. and Leder, P. (1979). *Cell*, 18: 1299-1302.
- Hamm, J., Kazmaier, M. and Mattaj, I.W. (1987). *EMBO J.* 6: 3479-3485.
- Hamm, J., Dathan, N.A. and Mattaj, I.W. (1989a). *Cell*, 59: 159-169.
- Hamm, J. and Mattaj, I.W. (1989b). *EMBO J.*, 8: 4179-4187.
- Hamm, J., Dathan, N.A., Scherly, D. and Mattaj, I.W. (1990). *EMBO J.*, 9: 1237-1244.
- Hardy, S.F., Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984). *Nature* 308: 375-377.
- Hartwell, L.H. (1967). *J. Bact.*, 93: 1662-1670.
- Hartwell, L.H., McLaughlin, C.S. and Warner, J.R. (1970). *Mol. Gen. Genet.*, 109: 42-56.
- Hashimoto, C. and Steitz, J.A. (1984). *Nucl. Acids Res.*, 12: 3282-3293.
- Hashimoto, C. and Steitz, J.A. (1986). *Cell*, 45: 581-591.
- Haynes, R.H. and Kunz, B.A. (1981) In: *The Molecular Biology of Saccharomyces*. (Strathern, Jones and Broach, Eds), pp. 371-414.
- Heinrichs, V., Bach, M. Winkelmann, G. and Lührmann, R. (1990). *Science*, 247: 69-72.
- Helfman, D.M., Roscigno, R.F., Mulligan, G.J., Finn, L.A. and Weber, K.S. (1990). *Genes Dev* 4: 98-110.
- Herman, R.C. (1989). *TIBS*, 14: 219-222.
- Hernandez, N., and Keller, W. (1983). *Cell*, 35: 89-99.
- Hicke, B.J., Christian, E.L. and Yarus, M. (1989). *EMBO J* 8: 3843-3851.
- Hitzeman, R.A., Hagie, F.E., Hayflick, J.S., Chen, C.Y., Seeburg, P.H. and Derynck, R. (1982). *Nucl. Acids Res.*, 10: 7791-7808.



- Hornig, H., Aebi, M. and Weissmann, C. (1986). *Nature* 324: 589-591.
- Hooper, A.K., Banks, F. and Evangelidis, V. (1978). *Cell*, 41: 211-219.
- Huang, M.T.F. and Gorman, C.M. (1990). *Nucl. Acids Res.*, 18: 937-946.
- Hutchison, T.H., Hartwell, L.H. and McLaughlin, C.S. (1969). *J. Bact.*, 99: 807-814.
- Igel, H.A. and Ares, M. Jr. (1988). *Nature* 334: 450-453.
- Inoue, K., Ohno, M., Sakamoto, H. and Shimura, Y. (1989). *Genes Dev.*, 3: 1472-1479.
- Inoue, K., Hoshijima, K., Sakamoto, H. and Shimura, Y. (1990). *Nature* 344: 461-463.
- Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983). *J. Bact.*, 153: 163-168.
- Jackson, S.P. (1987) Ph.D. Thesis, University of Edinburgh.
- Jackson, S.P., Lossky, M. and Beggs, J.D. (1988). *Mol. Cell. Biol.*, 8: 1067-1075.
- Jacob, M. and Gallinaro, H. (1989). *Nucl. Acids Res.*, 17: 2159
- Jaquier, A., Rodriguez, J.R. and Rosbash, M. (1985). *Cell*, 43: 423-430.
- Jaquier, A. and Rosbash, M. (1986). *Proc. Natl. Acad. Sci. USA*, 83: 5835-5839.
- Karn, J., Brenner, S. and Barnett, L. *Proc. Natl. Acad. Sci. USA*, 80: 4253-4257.
- Käuffer, N.F., Freid, H.M., Schwindinger, W.F., Jasin, M. and Warner, J.R. (1983). *Nucl. Acids Res.*, 11: 3123-3125.
- Kaufmann, R.F., Murtha, P. and Davies, M.V. (1987). *EMBO J.*, 6: 187-193.
- Kay, R.J., Russnak, R.H., Jones, D., Mathieas, C. and Candido E.P.M. (1987). *Nucl. Acids Res.* 15: 3723-3741.
- Keller, W. (1984). *Cell*, 39: 423-425.
- Keller, E.B. and Noon, W.A. (1984). *Proc. Natl. Acad. Sci. USA*, 81: 7317-7420.
- Khowry, G., Gruss, P., Dahr, R. and Lai, C. (1979). *Cell* 18: 85-92.
- Kim, C.H. and Warner, J.R. (1983). *Mol. Cell. Biol.*, 3: 457-465.
- Klinz, F-J. and Gallwitz, D. (1985). *Nucl. Acids Res.*, 13: 3791-2804.
- Knapp, G., Beckmann, J.S., Johnson, P.F., Fuhrman, S.A. and Abelson, J. (1978). *Cell*, 14: 221-236.

- Köhler, K. and Domdey, H. (1988). Nucl. Acids Res., 16: 9457-9475.
- Kole, R. and Weissman, S.M. (1982). Nucl. Acids Res., 10: 5429-5445.
- Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984). Cell, 38: 731-736.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1985). Nature 313: 552-557.
- Konarska, M.M. and Sharp, P.A. (1986). Cell, 46: 845-855.
- Konarska, M.M. and Sharp, P.A. (1987). Cell, 49: 763-774.
- Kozak, M. (1989a). Mol. Cell. Biol., 9: 5073-5080.
- Kozak, M. (1989b). Mol. Cell. Biol., 9: 5134-5142.
- Kozak, M. (1989c) J. Cell. Biol., 108: 229-241.
- Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984). Cell, 36: 993- 1005.
- Krainer, A.R., and Maniatis, T. (1985). Cell, 42: 725-736.
- Krainer, A.R. (1988). Nucl. Acids Res., 16: 9415-9429.
- Kramer, A., Keller, W., Appel, B. and Lührmann, R. (1984a). Cell, 38: 299- 307.
- Kramer, B., Kramer, W. and Fritz, H-J. (1984b). Cell, 38: 879-887.
- Kramer, A. (1988). Genes Dev., 2: 1155-1167.
- Kramer, A., Frick, M. and Keller, W. (1987). J. Biol. Chem., 262: 17630-17640.
- Kretzner, L., Rymond, B.C. and Rosbash, M. (1987). Cell, 50: 593-602.
- Lai, C. and Khoury, G. (1979). Proc. Natl. Acad. Sci. USA, 76: 71-75.
- Lambowitz, A.M. (1989). Cell, 56: 323-326.
- Lamond, A.I., Konarska, M.M. and Sharp, P.A. (1987). Genes Dev., 1: 532-543.
- Lamond, A.I., Konarska, M.M., Grabowski, P.J. and Sharp, P.A. (1988). Proc. Natl. Acad. Sci. USA, 85: 411-415.
- Lamond, A.I., Sproat, B., Ryder, U. and Hamm, J. (1989). Cell, 58: 383-390.
- Lamond, A.I. and Gibson, T.J. (1990). TIGS, 6: 145-149.
- Langford, C.J. and Gallwitz, D. (1983). Cell, 33: 519-527.
- Langford, C.J., Klinz, F., Donath, C. and Gallwitz, D. (1984). Cell, 36: 646-653.

- Last, R.L., Stavenhagen, J.B. and Woolford, J.L. Jr. (1984). *Mol. Cell. Biol.*, 4: 2396-2405.
- Last, R.L., and Woolford, J.L. (1986). *J. Cell Biol.*, 103: 2103-2112.
- Last, R.L., Maddock, J.R. and Woolford, J.L., Jr. (1987). *Genetics*, 117: 619-631.
- Lear, A.L., Eperon, L.P., Wheatley, I.M. and Eperon, I.C. (1990). *J. Mol. Biol.*, 211: 103-115.
- Lee, J. and Engelke, D.R. (1989). *Mol. Cell. Biol.*, 9: 2536-2543.
- Lee, M.G., Young, R.A. and Beggs, J.D. (1984). *EMBO J.*, 3: 2825-2830.
- Lee, M.G., Lane, D.P. and Beggs, J.D. (1986). *Yeast* 2: 59-67.
- Leer, R.J., van Raamsdonk-Duin, M.M.C., Hagendoorn, M.J.M., Mager, W.H. and Planta, R.J. (1984). *Nucl. Acids Res.*, 12: 6685-6700.
- Leff, S.E. and Rosenfeld, M.G. (1986). *Ann. Rev. Biochem.*, 55: 1091-1117.
- Legrain, P., and Rosbash, M. (1989). *Cell*, 57: 573-583.
- Lerner, M.R. and Steitz, J.A., (1979). *Proc. Natl. Acad. Sci. USA*, 76: 5495-5499.
- Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J.A. (1980). *Nature*, 283: 220-224.
- Lerner, E.A., Lerner, M.R., Janeway, C.A., Jr. and Steitz, J.A. (1981). *Proc. Natl. Acad. Sci. USA*, 78: 2737-2741.
- Liautard, J.P., Sti-Widada, J., Brunel, C. and Jeanteur, P. (1982). *J. Mol. Biol.*, 162: 623-643.
- Lin, R.J., Newman, A.J., Cheng, S-C. and Abelson, J. (1985). *J. Biol. Chem.*, 260: 14780-14792.
- Lin, R-J., Lustig, A.J. and Abelson, J. (1987). *Genes Dev.*, 1: 7-18.
- Lossky, M., Anderson, G.J., Jackson, S.P. and Beggs, J.D. (1987). *Cell*, 51: 1019-1026.
- Lund, E., Bostock, C.J. and Dahlberg, J.E. (1987). *Genes Dev.*, 1: 47-56.
- Lund, E. and Dahlberg, J.E. (1987). *Genes Dev.* 1: 39-46.
- Lührmann, R. (1988). In *Small Nuclear Ribonucleoprotein Particles*. Birnsteil, M.L. (ed). Springer-Verlag, pp. 71-99.
- Lustig, A.J., Lin, R-J. and Abelson, J. (1986). *Cell*, 47: 953-963.
- Lutz-Reyeremuth, C. and Keene, J.D. (1989). *Mol. Cell. Biol.*, 9: 2975-2982.

- McMaster, G.K. and Carmichael, G.G. (1977). *Proc. Natl. Acad. Sci. USA*, 74: 4835-4838.
- McPheeters, D.S., Fabrizio, P. and Abelson, J. (1989). *Genes Dev.*, 3: 2124-2136.
- Maniatis, T., Fritsch, T.F. and Sambrook, J. (1982). In: *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Maniatis, T. and Reed, R. (1987). *Nature* 325: 673-678.
- Mason, P.J. and Williams, J.G. (1985). In: *Nucleic Acid Hybridisation*. Hames, B.D. and Higgins, S.J. (eds), pp. 113-137.
- Mattaj, I.W. and De Robertis, E.M. (1985). *Cell*, 40: 111-118.
- Mattaj, I.W. (1986). *Cell*, 46: 905-911.
- Mattaj, I.W., Habets, W.J. and van Venrooije, W.J. (1986). *EMBO J.* 5: 997-1002.
- Mattaj, I.W. (1989). *Cell*, 57: 1-3.
- Mayrand, S.H., Pedersen, N. and Pederson, T. (1986). *Proc. Natl. Acad. Sci USA*, 83: 3718-3722.
- Messing, J., Crea, R. and Seeburg, P.H. (1981). *Nucl. Acids Res.*, 9: 309-321.
- Messing, J. (1983). In: *Methods in Enzymology* 101, Part C, pp. 20-78.
- Miller, A.M., (1984). *EMBO J.*, 3: 1061-1065.
- Miller, J.H. (1972). In: *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory.
- Mitra, G. and Warner, J.A. (1984). *J. Biol. Chem.*, 259: 9218-9224.
- Moore, C.L., Chen, J. and Whoriskey, J. (1988). *EMBO J.*, 7: 3159-3169.
- Mörl, M. and Schmelzer, C. (1990). *Cell*, 60: 629-636.
- Mount, S.M., (1982). *Nucl. Acids Res.*, 10: 459-472.
- Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983). *Cell*, 33: 509-518.
- Nagoshi, R.N. and Baker, B.S. (1990). *Genes Dev.*, 4: 89-97.
- Nelson, K.K. and Green, M.R. (1989). *Genes Dev.*, 3: 1562-1571.

- Nelson, K.K. and Green, M.R., (1988). *Genes Dev.*, 2: 319-329.
- Ner, S.S. and Smith, M. (1989). *Mol. Cell. Biol.*, 9: 4623-4630.
- Newman, A.J., Lin, R-J., Cheng, S-C. and Abelson, J. (1985). *Cell*, 42: 335-344.
- Newman, A. (1987). *EMBO J.*, 6: 3833-3839.
- Ng, R. and Abelson, J. (1980). *Proc. Natl. Acad. Sci. USA*, 77: 3912-3916.
- Noble, J.C.S., Prives, C. and Manley, J.L. (1988). *Genes Dev.*, 2: 1460-1475.
- Noble, J.C.S., Ge, H., Chaudhuri, M. and Manley, J.L. (1989). *Mol. Cell. Biol.*, 9: 2007-2017.
- Orgel, L.E. and Crick, F.H.C. (1980). *Nature*, 284: 604-607.
- Osheim, Y.N., Miller, O.L., Jr., and Beyer, A.L. (1985). *Cell*, 43: 143-151.
- Ottavio, L., Chang, C., Rizzo, M., Travali, S., Casadevall, C. and Baserga, R. (1990). *Mol. Cell. Biol.*, 10: 303-309.
- Padgett, R.A., Hardy, S.F. and Sharp, P.A. (1983a). *Proc. Natl. Acad. Sci. USA*, 80: 5230-5234.
- Padgett, R.A., Mount, S.M., Steitz, J.A. and Sharp, P.A. (1983b). *Cell*, 35: 101-107.
- Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984). *Science* 225: 898-903.
- Padgett, R.A., Konarska, M.M., Aebi, M., Hornig, H., Weissmann, C. and Sharp, P.A. (1985). *Proc. Natl. Acad. Sci. USA*, 82: 8349-8353.
- Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seilor, S. and Sharp, P.A. (1986). *Ann. Rev. Biochem.* 55: 1119-1150.
- Palfi, Z., Bach, M., Solymosy, F. and Lührmann, R. (1989). *Nucl. Acids Res.*, 17:
- Pan, Z-Q., Ge, H., Fu, X-Y., Manley, J.L. and Prives, C. (1989). *Nucl. Acids Res.*, 17: 6553-6568.
- Parent, A., Zeitlin, S. and Efstratiadis, A. (1987). *J. Biol. Chem.*, 262: 11284-11291.
- Parent, A., Wilson, R.C., Zeitlin, S. and Efstratiadis, A. (1989). *J. Mol. Biol.*, 209: 379-392.
- Parker, R. and Guthrie, C. (1985). *Cell*, 41: 107-118.
- Parker, R., Siliciano, P.G. and Guthrie, C. (1987). *Cell*, 49: 229-239.

- Parker, R., Simmons, T., Shuster, E.O., Siliciano, P.G. and Guthrie, C. (1988). *Mol. Cell. Biol.*, 8: 3150-3159.
- Patterson, B. and Guthrie, C. (1987). *Cell*, 49: 613-624.
- Patthy, L. (1987). *FEBS*, 214: 1-7.
- Patton, J.R., Patterson, R.J. and Pederson, T. (1987). *Mol. Cell. Biol.*, 7: 4030- 4037.
- Patton, J.R., Habets, W., van Venrooij, W.J. and Pederson, T. (1989). *Mol. Cell. Biol.*, 9: 3360-3368.
- Patzelt, E., Thalmann, E., Hartmuth, K., Blaas, D. and Kuechler, E. (1987). *Nucl. Acids Res.*, 15: 1387-1399.
- Peabody, D.S., Subramani, S. and Berg, P. (1986). *Mol. Cell. Biol.*, 6: 2704- 2711.
- Peabody, D.S. and Berg, P. (1986). *Mol. Cell. Biol.*, 6: 2695-2703.
- Pearson, N.J., Thorburn, P.C. and Haber, J.E. (1982). *Mol. Cell. Biol.*, 2: 571- 577.
- Perler, F., Efstratiadis, A., Lomdeico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980). *Cell*, 20: 555-566.
- Peterson-Björn, S., Soltyk, A., Beggs, J.D. and Friesen, J.D. (1989). *Mol. Cell. Biol.*, 9: 3698-3709.
- Pikielny, C.W., Teem, J.L. and Rosbash, M. (1983). *Cell*, 34: 395-403.
- Pikielny, C.W. and Rosbash, M. (1985). *Cell*, 41: 119-126.
- Pikielny, C.W., and Rosbash, M. (1986). *Cell*, 45: 869-877.
- Pikielny, C.W., Rymond, B.C. and Rosbash, M. (1986). *Nature*, 324: 341-345.
- Pikielny, C.W. Bindereif, A. and Green, M.R. (1989). *Genes Dev.*, 3: 479-487.
- Pinto, A.L. and Steitz, J.A. (1989). *Proc. Natl. Acad. Sci. USA*, 86: 8742-8746.
- Piper, P.W., Curran, B., Davies, M.W., Lockheart, A. and Reid, G. (1986). *Eur. J. Biochem.*, 161: 525-531.
- Query, C.C., Bentley, R.C. and Keene, J.D. (1989). *Mol. Cell. Biol.*, 9: 4872-4881.
- Reed, R. and Maniatis, T. (1985). *Cell*, 41: 95-105.

- Reed, R. and Maniatis, T. (1986). *Cell*, 46: 681-690.
- Reed, R., Griffith, J. and Maniatis, T. (1988). *Cell*, 53: 949-961.
- Reed, R. and Maniatis, T. (1988). *Genes Dev.*, 2: 1268-1276.
- Reed, R. (1989). *Genes Dev.*, 3: 2113-2123.
- Riedal, N., Wise, J-A., Swerdlow, H., Mak, A. and Guthrie, C. (1986). *Proc. Natl. Acad. Sci. USA*, 83: 8097-8101.
- Riedal, N., Wolin, S. and Guthrie, C. (1987). *Science* 235: 328-331.
- Robberson, B.L., Cote, G.J. and Berget, S.M. (1990). *Mol. Cell. Biol.*, 10: 84- 94.
- Rogers, J. and Wall, R. (1980). *Proc. Natl. Acad. Sci. USA*, 77: 1877-1879.
- Rodriguez, J.R., Pikielny, C.W. and Rosbash, M. (1984). *Cell*, 39: 603-610.
- Rosbash, M., Harris, P.K.W., Woolford, J.L. Jr. and Teem, J.L. (1981). *Cell*, 24: 679-686.
- Ruby, S.W. and Abelson, J. (1988). *Science*, 242: 1028-1036.
- Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984). *Cell*, 38: 317-331.
- Ruskin, B., Pikielny, C.W., Rosbash, M. and Green, M.R. (1986). *Proc. Natl. Acad. Sci. USA*, 83: 2022-2026.
- Ruskin, B. and Green, M.R. (1985a). *Cell* 43: 131-142.
- Ruskin, B. and Green, M.R. (1985b). *Science*, 229: 135-140.
- Ruskin, B., and Green, M.R. (1985c). *Nature* 317: 732-734.
- Ruskin, B., Greene, J.M. and Green, M.R. (1985). *Cell*, 41: 833-844.
- Ruskin, B., Zamore, P.D. and Green, M.R. (1988). *Cell*, 52: 207-219.
- Rymond, B.C. and Rosbash, M. (1988). *Genes Dev.*, 2: 428-439.
- Rymond, B.C. and Rosbash, M. (1985). *Nature*, 317: 735-737.
- Rymond, B.C., Torrey, D.D. and Rosbash, M. (1987). *Genes Dev.*, 1: 238-246.
- Rymond, B.C. and Rosbash, M. (1986). *EMBO J.*, 5: 3517-3523.
- Scazzocchio, C. (1989). *TIG*, 5: 168-172.

- Scherley, D., Beolens, W., van Venrooij, W.J., Dathan, N.A., Hamm, J. and Mattaj, I.W. (1989). EMBO J., 8: 4163-4170.
- Schmidt, F.J. (1985). Cell, 41: 339-340.
- Schulz, V.P. and Reznikoff, W.S. (1990). J. Mol. Biol., 211: 427-445.
- Seraphin, B., Kretzner, L. and Rosbash, M. (1988). EMBO J., 7: 2533-2538.
- Seraphin, B. and Rosbash, M. (1989a). Gene, 82: 145-151.
- Seraphin, B. and Rosbash, M. (1989b). Cell, 59: 349-358.
- Sharp, P.A. (1985). Cell, 42: 397-400.
- Sharp, P.A. (1987). Science, 235: 766-771.
- Shih, M-C., Heinrich, P. and Goodman, H.M. (1988). Science, 242: 1164-1166.
- Shiokawa, K. and Pogo, A.O. (1974). Proc. Natl. Acad. Sci. USA, 71: 2658-2662.
- Siliciano, P.G., Brow, D.A., Roiha, H. and Guthrie, C. (1987). Cell, 50: 585-592.
- Siliciano, P.G. and Guthrie, C. (1988). Genes Dev., 2: 1258-1267.
- Sillekens, P.T.G., Habets, W.J., Beijer, R.P. and van Venrooij, W.J. (1987). EMBO J., 6: 3841-3848.
- Sillekens, P.T.G., Beijer, R.P., Habets, W.J. and van Venrooij, W.J. (1988). Nucl. Acids Res., 16: 8307-8321.
- Smith, J.H. and Eliceiri, G.L. (1983). J. Biol. Chem., 258: 4636-4640.
- Smith, C.W.J., Patton, J.G. and Nadal-Ginard, B. (1989a). Ann. Rev. Genet., 23: 527-577.
- Smith, C.W.J., Porro, E.B., Patton, J.G. and Nadal-Ginard, B. (1989b). Nature 342: 243-248.
- Sogin, M.L., Ingold, A., Karlock, M., Nielsen, H. and Engberg, J. (1986). EMBO J., 5: 3625-3630.
- Solnick, D. (1985). Cell, 43: 667-676.
- Sonenberg, N. and Pelletier, J. (1989). Bio-Essays 11: 128-132.
- Sosnowski, B.A. (1989). Cell, 58: 449-459.



- Stanford, D.R., Kehl, M., Perry, C.A., Holicky, E.L., Harvey, S.E., Rohleder, A.M., Rehder, K., Jr., Lührmann, R. and Weiben, E.D. (1988). Nucl. Acids Res., 16: 10593-10605.
- Steitz, J.A. (1988). Sci. Amer., 258: 36-41.
- St John, T.P. and Davis, R.W. (1981). J. Mol. Biol., 152: 285-315.
- Struhl, K. and Davis, R.W. (1981). J. Mol. Biol., 152: 535-552.
- Südhof, T.C., Goldstein, J.L., Brown, M.S. and Russell, D.W. (1985a) Science 228: 815-822.
- Südhof, T.C., Russell, D.W., Goldstein, J.L., Brown, M.S., Sanchez-Pescador, R. and Bell, G.I. (1985b). Science 288: 893-895.
- Suggs, S.V., Hirose, T., Miyaki, T., Kawashima, E.H., Johnson, M.J., Itakura, K. and Wallace, R.B. (1981). In: Developmental Biology Using Purified Genes. Brow, D.D. and Fox, C.F. (eds), pp. 683-693.
- Surowy, C.S., van Santen, V.L., Schreib-Wixted, S.M. and Spritz, R.A. (1989). Mol. Cell. Biol., 9: 4179-4186.
- Swanson, M.S. and Dreyfuss, G. (1988). EMBO J., 7: 3519-3529.
- Swida, U., Thüroff, E. and Käufer, N.F. (1986). Mol. Gen. Genet., 203: 300-304.
- Tani, T. and Ohshima, Y. (1989). Nature 337: 87-90.
- Tatei, K., Teakemura, K., Tanaka, H., Mosaki, T. and Ohshima, Y. (1984). J. Biol. Chem., 259: 11667-11674.
- Tazi, J., Alibert, C., Temsamani, J., Reveillaud, I., Cathala, G., Brunel, C. and Jeanteur, P. (1986). Cell, 47: 755-766.
- Tazi, J., Temsamani, J., Alibert, C., Rhead, W., Khellil, S., Cathala, G., Brunel, C. and Jeanteur, P. (1989). Nucl. Acids Res., 17: 5223-5243.
- Teem, J.L. and Rosbash, M. (1983). Proc. Natl. Acad. Sci. USA, 80: 4403-4407.
- Teem, J., Abovich, N., Käufer, N., Schwindinger, W., Warner, J., Levy, A., Woolford, J., Leer, R., Van Raamschön, R., Duin, M., Mager, W., Planta, R., Schultz, L., Friesen, J., Freid, H. and Rosbash, M. (1984). Nucl. Acids Res., 12: 8295-8312.

- Temsamani, J., Alibert, C., Tazi, J., Rucheton, M., Capony, J-P., Jeanteur, P., Cathala, G. and Brunel, C. (1989). *J. Mol. Biol.*, 206: 439-449.
- Thompson-Jäger, S. and Domdey, H. (1987). *Mol. Cell. Biol.*, 7: 4010-4016.
- Thompson-Jäger, S. and Domdey, H. (1990). *Curr. Genet.*, 17: 269-273.
- Tollervey, D. (1987). *J. Mol. Biol.*, 196: 355-361.
- Tollervey, D. and Mattaj, I.W. (1987). *EMBO J.*, 6: 469-476.
- Tournier-Lasserre, E., Odenwald, W.F., Garbern, J., Trojanowski, J. and Lazzarini, R.A. (1989). *Mol. Cell. Biol.*, 9: 2273-2278.
- Traglia, H.M., Atkinson, N.S. and Hopper, A.K. (1989). *Mol. Cell. Biol.*, 9: 2989-2999.
- Treisman, R., Novak, U., Favalaro, J. and Kamen, R. (1981). *Nature*, 292: 595-600.
- Treisman, R., Proudfoot, N.J., Shander, M. and Maniatis, T. (1982). *Cell*, 29: 903-911.
- Treisman, R., Orkin, S.H. and Maniatis, T. (1983). *Nature* 302: 591-596.
- Tsai, A.Y.M., Streuli, M. and Saitio, H. (1989). *Mol. Cell. Biol.*, 9: 4550-4555.
- Turnbull-Ross, A.D., Else, A.J. and Eperon, I.C. (1988). *Nucl. Acids Res.*, 16: 395-411.
- Ulfendahl, P.J., Kreivi, J-P. and Akusjarvi, G. (1989). *Nucl. Acids Res.*, 17: 925-938.
- Van Dam, A., Winkel, I., Zijlstra-Baalbergen, J., Smeenk, R. and Theo Cuipers, H. (1989). *EMBO J.*, 8: 3853-3860.
- van Santen, V.L. and Spritz, R.A. (1985). *Proc. Natl. Acad. Sci. USA*, 82: 2885-2889.
- Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J. and Guthrie, C. (1986). *EMBO J.*, 5: 1683-1695.
- Vijayraghavan, U., Company, M. and Abelson, J. (1989). *Genes Dev.*, 3: 1206-1216.
- Vijayraghavan, U. and Abelson, J. (1990). *Mol. Cell. Biol.*, 10: 324-332.
- Wallace, J.C. and Edmonds, M. (1983). *Proc. Natl. Acad. Sci. USA*, 80: 950-954.
- Warner, (1987). *Genes Dev.*, 1 1-3.
- Weingartner, B. and Keller, W. (1981). *Proc. Natl. Acad. Sci. USA*, 78: 4092-4096.
- Weissman, C. (1984). *Nature*, 311: 103-104.

- Whittaker, E., Lossky, M. and Beggs, J.D. (1990). *Proc. Natl. Acad. Sci. USA*, 87: 2216-2219.
- Wickens, M.P., Woo, S., O'Malley, B.W. and Gurdon, J.B. (1980). *Nature*, 285: 628-634.
- Wieringa, B., Meyer, F., Reiser, J. and Weissmann, C. (1983). *Nature*, 301: 38-43.
- Wieringa, B., Hofer, E. and Weissmann, C. (1984). *Cell*, 37: 915-925.
- Wilson, W., Braddock, M., Adams, S.E., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988). *Cell*, 55: 1159-1169.
- Winkelmann, G., Bach, M. and Lührmann, R. (1989). *EMBO J.*, 8: 3105-3112.
- Woodson, S.A. and Cech, T.R. (1989). *Cell*, 57: 335-345.
- Wu, J. and Manley, J.L. (1989). *Genes Dev.*, 3: 1553-1561.
- Xu, Y., Petersen-Björn, S. and Friesen, J.D. (1990). *Mol. Cell. Biol.*, 10: 1217-1225.
- Yee-Yuo, C. and Weiner, A.M. (1989a). *Mol. Cell. Biol.*, 9: 3429-3437.
- Yee-Yuo, C. and Weiner, A.M. (1989b). *Genes Dev.*, 3: 697-707.
- Zapp, M.L. and Berget, S.M. (1989). *Nucl. Acids Res.*, 17: 2655-2674.
- Zeitlin, S. and Efstratiadis, E. (1984). *Cell*, 39: 589-602.
- Zeitlin, S., Parent, A., Silverstein, S. and Efstratiadis, E. (1987). *Mol. Cell. Biol.*, 7: 111-120.
- Zhuang, Y. and Weiner, A.M. (1986). *Cell*, 46: 827-835.
- Zhuang, Y. and Weiner, A.M. (1989). *Genes Dev.*, 3: 1545-1552.
- Zhuang, Y., Goldstein, A.M. and Weiner, A.M. (1989). *Proc. Natl. Acad. Sci. USA*, 86: 2752-2756.
- Zillman, M., Rose, S. and Berget, M. (1987). *Mol. Cell. Biol.*, 7: 2877-2883.
- Zillmann, M., Zapp, M.L. and Berget, S.M. (1988). *Mol. Cell. Biol.*, 8: 814-821.
- Zoller, M.J. and Smith, M. (1984). *DNA*, 3: 479-488.

## APPENDIX

### COMMITMENT OF A PRE-mRNA TRANSCRIPT TO SPLICING

#### INTRODUCTION

Previous *in vitro* studies have demonstrated that pre-mRNA is committed to splicing prior to the formation of the U2 snRNP-containing prespliceosome (Seraphin *et al.*, 1988; Legrain *et al.*, 1988; Seraphin and Rosbash, 1989b). A "commitment" complex has been identified and its formation (which requires little or no ATP) relies upon the presence on the transcript of the conserved 5' splice-site and branch-point sequences (Seraphin and Rosbash, 1989b). This complex contains the U1 snRNP. A rather subtle *in vivo* approach was used to identify *cis*- and *trans*-acting factors which might interfere with this commitment stage (Legrain and Rosbash, 1989). Based on the observations made by this approach, the authors proposed a model for the commitment of pre-mRNA to splicing. This Appendix presents experiments designed to complement and extend such an approach. Please refer to Legrain and Rosbash (1989) for a full description of the experiments undertaken. A brief account of the pre-mRNA commitment model is given below.

Newly transcribed pre-mRNA is bound by a commitment factor(s) (at least one of which is the U1 snRNP) which determines that the transcript enters the splicing pathway. The binding of this commitment factor(s) inhibits the transport of pre-mRNA to the cytoplasm, retaining it in the nucleus until the intron has been removed. In this model intron-less transcripts are not retained in the nucleus because they lack the conserved 5' splice site and branchpoint sequences which are required for the formation of the commitment complex.

The rationale for the experiments used to establish the commitment model is that intron-containing transcripts may be retarded in the nucleus by the splicing machinery in contrast to their intron-less counterparts. If these respective transcripts encode  $\beta$ -galactosidase fusion-proteins, then more activity should be measureable from the intron-less transcripts because at any one time more of them are present in the cytoplasm (even though the rate of transcription is the same for intron-less and intron-containing). Therefore pJBM-5 should

produce more  $\beta$ -galactosidase activity than pJBM-1 (see Table below). The difference in activities arising from pJBM-1 and pJBM-1/s (if any) should be due to pre-mRNA transcripts (generated from pJBM-1) which leave the nucleus before being committed to splicing, and this should be equivalent to the activity produced by pJBM-4.

#### ANALYSIS OF PRE-mRNA COMMITMENT UTILISING pJBM-1 TO -5

In an attempt to repeat part of the analyses carried out by Legrain and Rosbash (1989), DBY745 (the same strain as used by these authors) was transformed with pJBM-1 to pJBM-5 (see Chapter 3.4) and assayed for  $\beta$ -galactosidase activity as described in Chapter 2.2.9. As previously indicated the pJBM-4 fusion-transcript containing the *MATa1* intron is spliced with roughly 80% efficiency (see Chapter 4.1.1). This contrasts with the transcript utilised in the study of Legrain and Rosbash which is spliced with 20% efficiency. It was considered that a repeat of the experiments of Legrain and Rosbash, incorporating a "normal" yeast intron, would indicate whether the results observed were general for yeast intron-containing transcripts. In the following sections cultures of transformed strains are described thus; the name of the strain e.g. DBY745, followed by the name of the plasmid residing in it, in brackets, e.g. (pJBM-5). A table indicating the equivalent plasmids (in terms of their  $\beta$ -galactosidase producing ability) used in the study of Legrain and Rosbash and in this study, is presented below.

<u>Plasmid Name</u> (Legrain and Rosbash)	<u>Plasmid Name</u> (This Study)	<u><math>\beta</math>-galactosidase</u>	
		<u>Pre-mRNA</u>	<u>mRNA</u>
pLGSD5	pJBM-5		+
PLint	pJBM-1	+	+
Acc <sup>o</sup>	pJBM-1/s	-	+
Nde <sup>o</sup> Acc <sup>o</sup>	pJBM-4	+	-
Sty <sup>o</sup>	pJBM-4/s	-	-

#### Comparison of pJBM-1, -1/s and -5

Figs. 6.1 and 6.2 compares the  $\beta$ -galactosidase activities exhibited by DBY745 (pJBM-1), (pJBM-1/s) and (pJBM-5) when incubated at 23°C and 36°C respectively. In all experiments for each plasmid, two independent transformants (colony purified 3 times) were tested and each was assayed in duplicate. The results presented are the average values for each plasmid i.e. the results of the two transformants (of a specific plasmid) were combined and averaged.

On incubation at 23°C and 36°C, DBY745 (pJBM-1) produces greater activity when compared to DBY745 (pJBM-5). Such results are in complete contrast to those obtained by Legrain and Rosbash. At 23°C, cultures of DBY745 (pJBM-1/s) showed a similar level of activity to those of DBY745 (pJBM-5) (at least for the first 4 hours). However activity arising from pJBM-1/s is greater than that arising from pJBM-5 when the cultures are incubated 36°C. Comparison of DBY745 (pJBM-1) and (pJBM-1/s) reveals that the activity of each is greater at 36°C relative to 23°C. In contrast DBY745 (pJBM-5) showed similar  $\beta$ -galactosidase activities at both temperatures (see Figs. 4.2, 6.1 and 6.2). It appears that splicing may be more efficient when cultures are incubated at 36°C (see Chapter 4.1 and 4.2). Thus increased levels of mRNA caused by this increase in splicing efficiency may explain the increase in  $\beta$ -galactosidase activity of DBY745 (pJBM-1) and DBY745 (pJBM-1/s) at 36°C (see Table above).

Cultures of DBY745 (pJBM-1/s) incubated at 23°C or 36°C exhibit roughly 66% of the activity of DBY745 (pJBM-1). According to the observations of Legrain and Rosbash the difference in activities arising from pJBM-1 and -1/s should equal the pre-mRNA-encoded activity arising from pJBM-4 (see the difference between PLint and Acc') which therefore should produce 34% of the activity of DBY745 (pJBM-1). However activity exhibited by DBY745 (pJBM-4), on incubation at 23°C, is only 1.57% (all time-points averaged) of the activity of DBY745 (pJBM-1) (compare Figs. 4.3, 6.4 and 6.1). This may mean that a large proportion of the  $\beta$ -galactosidase activity generated from pJBM-1 is not encoded by the pre-mRNA. However other interpretations are feasible e.g. the pJBM-1/s *MATa1-lacZ* fusion transcript may be poorly spliced, and/or the gene otherwise poorly expressed resulting in its decreased expression relative to pJBM-1 (also see below).

Variations in  $\beta$ -galactosidase activities could be a result of differences in the stability of the pre-mRNA and/or  $\beta$ -galactosidase fusion proteins encoded by pJBM-1 to pJBM-5. Instability could be a consequence of the specific mutations introduced into the plasmid constructs and/or the removal of the intron sequences due to splicing (see Chapter 3 and Discussion). Although such differences were not detected in the study of Legrain and Rosbash (it should be noted that reference is made only to those RNAs that could encode  $\beta$ -galactosidase, presumably others would be unstable in the cytoplasm due to the absence of "protecting" ribosomes), the possibility of differential instability of RNA and/or protein cannot be ruled out.

Another possible reason for the anomalous  $\beta$ -galactosidase activities is the presence of internal AUG codons on the fusion-transcripts generated from pJBM-1 and -1/s (see Chapter 3.3.3.2). Initiation of translation at these codons on a transcript not normally able to encode a  $\beta$ -galactosidase fusion protein, could theoretically lead to the production of  $\beta$ -galactosidase activity. The plasmid pJBM-2 encodes a *MATa1-lacZ* fusion-transcript that contains three internal AUG codons which could be possible sites of initiation (NB. it differed from the fusion-transcript generated from pJBM-4 only in this respect). It was hypothesised that the activity exhibited by DBY745 (pJBM-2) could possibly produce the 34% difference in activities

**FIGURE 6.1: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES PRODUCED BY pJBM-5, -1 and -1/s IN DBY745 INCUBATED AT 23°C**

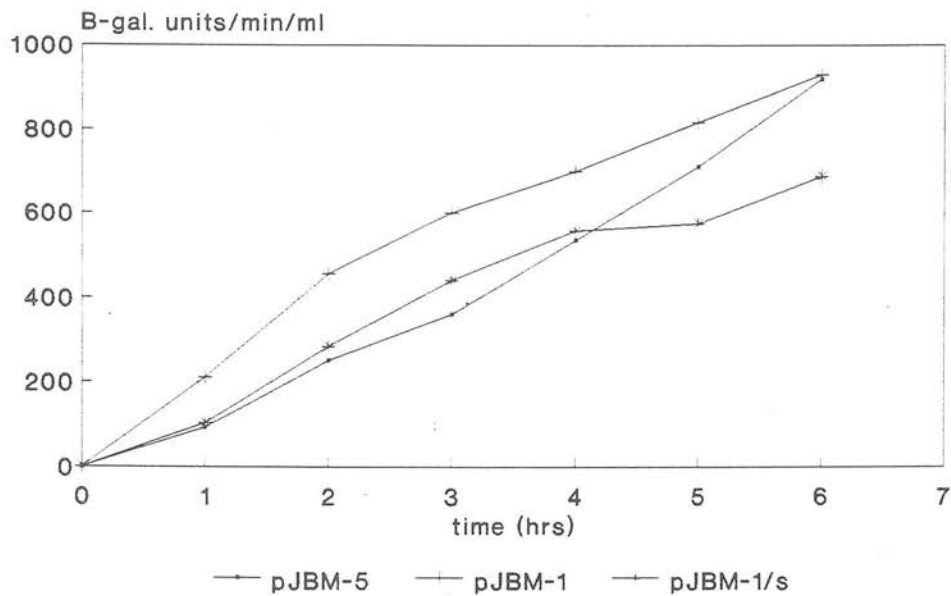
Cultures were grown at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 prior to dilution and galactose induction, after which incubation was continued at 23°C (see Chapter 2.2.9). Culture-samples (0.5ml) were taken at hourly intervals after galactose induction and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately after the addition of galactose, the so-called time-zero (see Chapter 2.2.9.1) and therefore 0 hours induction has an activity value of 0 (see Fig. 6.1).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

**FIGURE 6.2: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES PRODUCED BY pJBM-5, -1 AND -1/s IN DBY745 INCUBATED AT 36°C**

All experimental details were as for Fig. 6.1 except that the cultures were initially grown at 23°C to an OD<sub>600nm</sub> of 0.1-0.4 then heat-shocked and incubated at 36°C. Fifteen minutes after the heat-shock at 36°C galactose was added to a final concentration of 2% and this is time-zero (see Chapter 2.2.9.1).

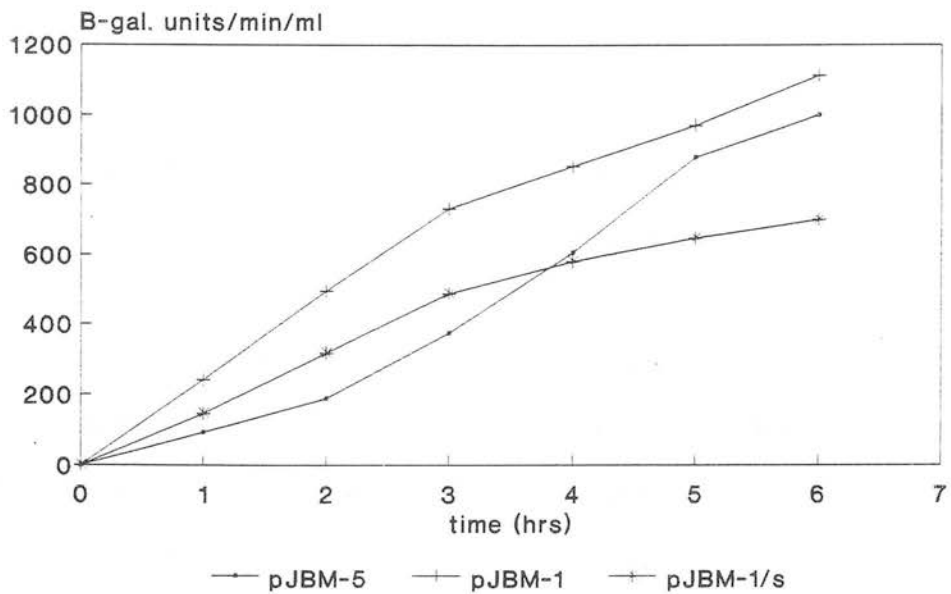


**FIG. 6.1**  
pJBM-5, -1 and -1/s at 23°C



B-gal. activities in DBY745

**FIG. 6.2**  
pJBM-5, -1 and -1/s at 36°C



B-gal. activities in DBY745

observed between DBY745 (pJBM-1) and (pJBM-1/s) (see above). Figs. 6.3 and 6.4 compare the activities of DBY745 (pJBM-2) and (pJBM-4) when incubated at 23°C and 36°C respectively. The activities of DBY745 (pJBM-2) incubated at both temperatures are roughly 3 times those of DBY745 (pJBM-4) (4 hour time-point onwards). This suggests that there is a substantial level of internal initiation of translation, however the activity exhibited by DBY745 (pJBM-2) is only approximately 4% of that produced by DBY745 (pJBM-1) and not the 34% difference in activity observed between DBY745 (pJBM-1) and (pJBM-1/s). It must be remembered that pJBM-2 and 4 differ only in the presence of these three AUGs.

To determine if this effect was a strain dependent phenomenon a similar analysis was carried out in the yeast strain SI50-2B. Fig. 6.5 compares the activities generated in cultures of SI50-2B (pJBM-2) incubated at 30°C with cultures of SI50-2B (pJBM-4). Internal initiation of translation may have been a consequence of incubating the cultures at sub-optimal temperatures i.e. 23°C and 36°C, therefore in this experiment the cultures were incubated at the optimal temperature for yeast growth. Activity arising from pJBM-2 is roughly 3-5 times greater than that generated by pJBM-4 (2-4 hour time-points), suggesting that this effect is not strain dependent.

It was considered that there had to be another contributing factor(s) to account for the 34% increase in activity arising from pJBM-1 compared to pJBM-1/s. The only difference between pJBM-1 and -1/s (apart from the stop codon in the intron) is the absence in pJBM-1/s of an AUG at the 5' end of the *lacZ* gene (see Chapter 3.3.1.3) i.e. pJBM-1/s has only two of the internal initiation codons. Thus the greater activity observed for cultures of DBY745 (pJBM-1) i.e. above what would be predicted according to the observations of Legrain and Rosbash, could be a result of initiation at this AUG. Indeed the two internal AUG codons present in pJBM-1/s may explain the relative increase in activity of that construct compared to pJBM-5, when a comparison between the results of Legrain and Rosbash and those presented here is made (pJBM-5 does not contain the three internal AUG codons). Because of this possible phenomenon of internal initiation of translation no conclusions regarding the process of pre-mRNA commitment can be drawn from the comparative analysis of pJBM-1, -1/s and -5.

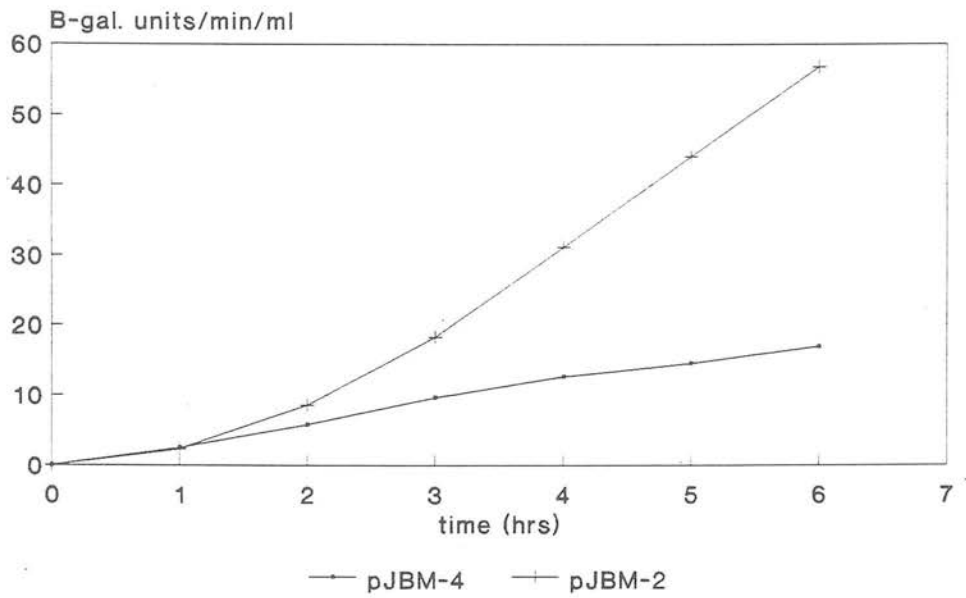
**FIGURE 6.3: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES PRODUCED BY  
pJBM-4 AND pJBM-2 IN DBY745 INCUBATED AT 23°C**

All experimental details were as for Fig. 6.1 except that 0.5-2ml culture-samples were taken.

**FIGURE 6.4: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES PRODUCED BY  
pJBM-4 AND pJBM-2 IN DBY745 INCUBATED AT 36°C**

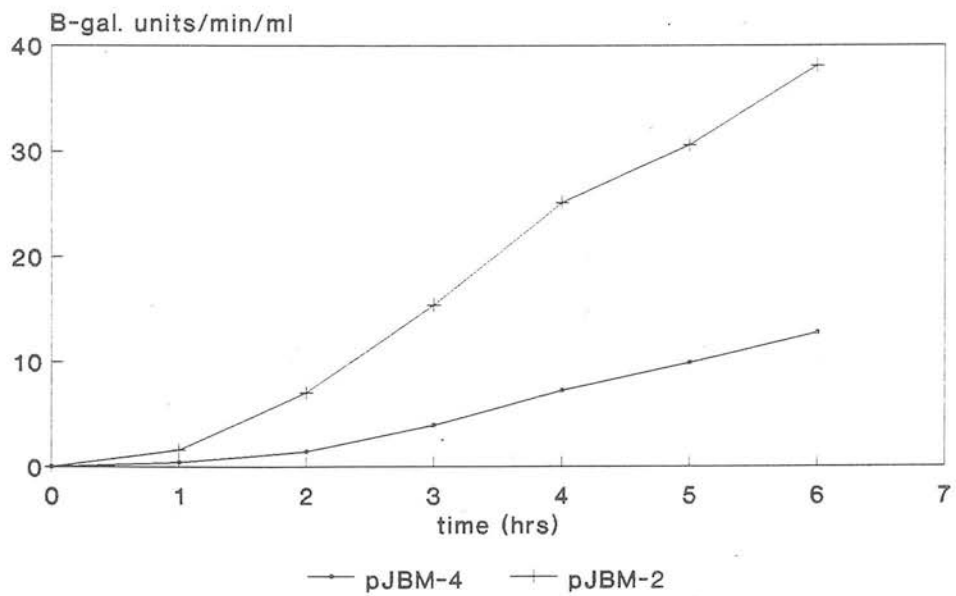
All experimental details were as for Fig. 6.2 except that 0.5-2ml culture-samples were taken.

**FIG. 6.3**  
pJBM-4 and pJBM-2 at 23°C



B-gal. activities in DBY745

**FIG. 6.4**  
pJBM-4 and pJBM-2 at 36°C

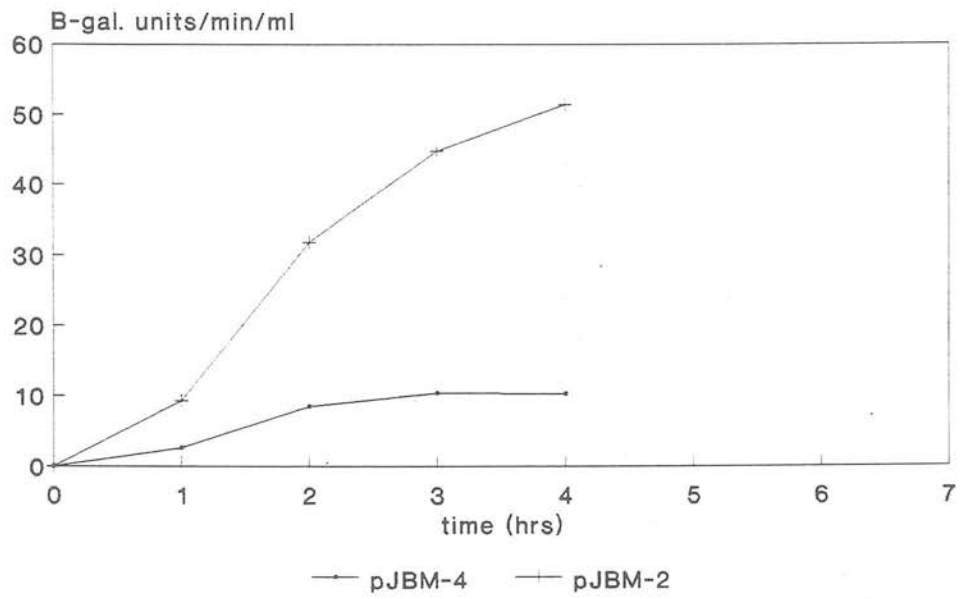


B-gal. activities in DBY745

**FIGURE 6.5: COMPARISON OF  $\beta$ -GALACTOSIDASE ACTIVITIES PRODUCED BY pJBM-4 AND pJBM-2 IN S150-2B INCUBATED AT 30°C**

Cultures were grown at 30°C to an OD<sub>600nm</sub> of 0.1-0.4 prior to galactose induction, after which incubation was continued at 30°C (see Chapter 2.2.9). Culture-samples (0.5-2ml) were taken at hourly intervals after galactose induction, and assayed for  $\beta$ -galactosidase activity (see Chapter 2.2.9.2). The results presented represent the average values obtained from two transformants, both of which were assayed in duplicate. All activities were corrected for background activity present immediately after the addition of galactose, the so-called time-zero (see Chapter 2.2.9.1) and therefore 0 hour induction has an activity value of 0 (see Fig. 6.5).  $\beta$ -galactosidase units were calculated as described in Chapter 2.2.9.2.

**FIG. 6.5**  
pJBM-4 and pJBM-2 at 30°C



B-gal. activities in S150-2B

Cultures of DBY745 (pJBM-4) incubated at 23°C exhibit roughly 2.3% of the activity of DBY745 (pJBM-5) (see Chapter 4, Figs. 4.2 and 4.3). Neither plasmid contains the three internal AUGs described above (see Chapter 3.3.1.3 and 3.3.3.2). This result is in close agreement with that observed for cultures of DBY745 (Nde°Acc°) and (pLGSD5). However the introns used in either study are spliced from their respective transcripts with different efficiencies (see above and Chapter 4). This means that although at any one time cells transformed with Nde°Acc° contain vastly more unspliced pre-mRNA, roughly the same amount of pre-mRNA leaves the nucleus as that observed for cells transformed with pJBM-4 (N.B. that both analyses were performed in DBY745). These results suggest that transcripts with completely contrasting splicing efficiencies are committed to splicing with a similar efficiency.

One problem with this analysis is that the activity of DBY745 (pJBM-4/s) at 23°C accounts for roughly 45% of the activity of DBY745 (pJBM-4). This is not the case with DBY745 (Sty°) and (Nde°Acc°). If the activity arising from pJBM-4/s results from read-through of the stop-codon in the pre-mRNA (as suspected), this means that the pre-mRNA generated from pJBM-4 has still exited the nucleus and that the percentage activity measured is indeed 2.3% of that arising from pJBM-5. However if it is due to alternative splicing and/or frameshifting then it is an activity encoded by mRNA and therefore has to be subtracted from pJBM-4, leaving roughly 1.27% compared to the 2.3% measured (see Chapter 4.1.2.1 for details of the pJBM-4/s percentage activity of pJBM-4). If the latter scenario is correct then there may be a possible explanation for the slight comparative difference in activity exhibited by DBY745 (pJBM-4) and (Nde°Acc°) (1.27% compared to roughly 3%). The higher ratio of activity observed for DBY745 (Nde°Acc°) in comparison to DBY745 (pJBM-4) may be a consequence of the fact that the intron present on the Nde°Acc° generated fusion-transcript (the transcript which encodes the  $\beta$ -galactosidase fusion protein) is spliced inefficiently and thus may produce a modest titration of commitment factors. This would lead to an exit of more pre-mRNA from the nucleus and thus effect the difference in ratios between pJBM-4 and Nde°Acc° in comparison to their respective controls.

## DISCUSSION

In an attempt to repeat part of the analyses carried out by Legrain and Rosbash, the  $\beta$ -galactosidase activity profiles exhibited by cultures of DBY745 (pJBM-1) to (pJBM-5) were obtained. Interpretation of the results was complicated by the demonstration that there appeared to be internal initiation of translation occurring from at least one of three internal AUG codons present on the transcripts generated by some of the plasmids. However anomalous results could also have been a consequence of RNA and/or protein instability.

In contrast to the situation in higher eukaryotes, initiation of translation in yeast is largely independent of sequence context surrounding the initiating AUG (Cigan and Donahue, 1987; Kozak, 1989a, 1989c). Recent reports suggest that nucleotides at positions (-1) and (-3), with respect to the AUG, are important in yeast in determining the efficiency of initiation; with an A preferred at (-3) (Baim and Sherman, 1988). However although a consensus sequence round the AUG can be derived for yeast, it appears that context effects are not extensive. In fact some yeast genes have initiating AUGs situated in regions that clearly differ from the consensus sequence e.g. *PPR1* and *PET3* (Brown, 1989 and see below). In a study of initiation sites of yeast gene transcripts it was concluded that 95% initiate translation at the first AUG. *MATa1* was included in this analysis and is thought to initiate at the first AUG on the transcript (Cigan and Donahue, 1987). The sequence context of the internal AUGs (which could give rise to  $\beta$ -galactosidase activity) present on the transcripts encoding the *MATa1-lacZ* fusion proteins, are all different and thus may affect the relative choice of initiation sites. Initiation at non-AUG codons has been detected in yeast transcripts but appears to be relatively inefficient (Brown, 1989). Initiation of this kind (if any) would probably be similar on fusion-transcripts generated by the constructs pJBM-1 to pJBM-5, unless sequence differences which led to secondary structure variations had an effect on their use. Therefore this would not be expected to present a problem for this study.

Initiation of translation in yeast as in higher eukaryotes can be influenced by the presence of secondary structure in the vicinity of the first AUG (Baim and Sherman, 1988;



Kozak, 1989b). This regulation is dependent on the position of RNA hairpins that appear to inhibit translation when upstream of the AUG, yet enhance it when further downstream, presumably by slowing up the 40S ribosomal initiation complex (Kozak, 1989b). The presence of RNA secondary structures may be a possible explanation for the use of alternative AUGs. The normal size of the *MATa1* transcript leader sequence is 11 nucleotides but pJBM-1 to -5 transcripts contain the *GAL1* transcript leader which is roughly 60 nucleotides long. This RNA extension at the 5' end of the *MATa1-lacZ* transcript may contribute secondary structures which may exclude the initiating AUG; however the "scanning" model of translation initiation would predict that in this case downstream AUGs would not be utilised (Kozak, 1989b). It is not known whether the ribosomal initiation complex can by-pass secondary structure which shields the first AUG to initiate at an internal AUG, without firstly unwinding the RNA and presumably recognising the first AUG, as predicted by the "scanning" model (Kozak, 1989b). This situation is further complicated by the fact that splicing produces different sequences with potentially different secondary structures; indeed the introduction of single mutations may induce similar effects. These changes may produce subtle or drastic effects on the initiation process which may enable the use of more than one site on a given transcript. In this respect it has been demonstrated that changes in the secondary structure at the 5' end of the *E.coli lacZ* transcript leader sequence can produce different  $\beta$ -galactosidase activities by either aiding or obstructing initiation of translation (Schulz and Reznikoff, 1990). It is not unreasonable to speculate that secondary structures arising in the *GAL1* and *MATa1* sequences (spliced and/or unspliced) may thus affect translation of the *MATa1-lacZ* fusion transcript.

In higher eukaryotic systems it has been shown that ribosomes can scan past upstream AUGs to initiate translation at downstream AUG codons (Kozak, 1989c). This process is known as "leaky" scanning and is dependent on the upstream AUGs being situated in a poor sequence context for initiation. Also of importance is the observation that ribosomes can reinitiate translation at a downstream site on the same transcript as long as the previous translational event has been terminated (Peabody *et al.*, 1986; Peabody and Berg, 1986). Efficient reinitiation relies on the site of termination being relatively close (roughly 100nts) to the reinitiation site

and indeed termination can be either side of the reinitiation codon, suggesting that a ribosome can scan the transcript in both directions. Another interesting possibility is that eukaryotic ribosomes can bind directly to internal sites on the transcript rather than scan along to them from the 5' end. Such a phenomenon has been reported for a few animal viruses (Herman, 1989; Sonenberg and Pelletier, 1989), however a need for caution in the interpretation of these results has been argued vigorously (Kozak, 1989c).

There are several examples of yeast transcripts that contain AUG codons in the leader sequence upstream of the main initiating AUG, e.g. *CPA1*, *GCN4*, *PET3*, *PPR1*, *SUC2* etc. (Brown, 1989). In some of these cases e.g. *CPA1* and *GCN4*, it is postulated that *trans*-acting factors may regulate the relative use of these codons. The mechanism by which the initiating complex recognises the downstream AUG is unknown. Indeed it may be a combination of several considerations i.e. sequence contexts, secondary structure and *trans*-acting factors. The upstream AUGs can be in variable sequence contexts, at different positions relative to the initiating AUG, be in or out of frame with the main AUG and with or without downstream termination signals. It is obvious that regulation of gene expression at the level of translation is complex and may be transcript dependent.

Another possibility which may lead to production of  $\beta$ -galactosidase activity from the mRNA encoded by pJBM-2 and not pJBM-4, is that translational frame-shifting on transcripts generated from pJBM-2 is greater than on those generated from pJBM-4. If this happened on the pre-mRNA *MATa1-lacZ* fusion transcripts generated by these constructs then no activity would be encoded (pre-mRNA is already in-frame). However it could occur on the mRNA which is out of frame and lead to increased  $\beta$ -galactosidase activity. Sequence differences between the *MATa1-lacZ* fusion transcripts generated from pJBM-2 and 4 (albeit very minor) could account for different rates of frame-shifting. This would also apply to the other constructs. Once again different sequence contexts may produce variations in the relative efficiencies of frame-shifting and result in different  $\beta$ -galactosidase activities. Yeast does in fact appear to exhibit translational frame-shifting at a level of 5-10% of normal translation (Wilson *et al.*, 1988). However this relied on the presence of six U nucleotides on the transcript immediately

upstream of the frame-shift site and it was suggested that the ribosomes fail to translocate properly due to the uniformity of this sequence. Indeed mutations in this sequence which replaced certain U nucleotides inhibited the frame-shift. Secondary structures in the RNA may also affect frame-shifting, though this was excluded in the study of Wilson *et al.* (Brown, 1989; Wilson, *et al.*, 1988 and references therein). There does not appear to be anything analogous to the U rich sequence described, in the *GAL1* or *MATa1* regions of pJBM-1 to -5 and thus frame-shifting may not occur. However the mechanism(s) responsible for this process has not been elucidated and therefore other unidentified factors present on these *MATa1-lacZ* fusion transcripts may permit frame-shifting to occur.

It is apparent that any sequence changes and RNA processing steps (including RNA editing, changes in termination and/or polyadenylation) which could affect secondary structure and/or consensus sequences required for translation initiation, have the potential to influence both translational initiation and possibly the rates of translation between different transcripts. Thus the interpretation of results based on translation which are extrapolated to accommodate changes in the RNA itself must take account of these possible problems.

Future work on the commitment of pre-mRNA to splicing will involve the construction of plasmids with the coding capacity of pJBM-1 and -1/s but without the internal AUGs. These will be derived from pJBM-4 and 4/s respectively. In this way it is hoped to generate transcripts that are more comparable with each other and to possibly control for the potential problems indicated above (it should be noted that pJBM-5 is derived from pJBM-4). However because the analysis relies on splicing, the potential secondary structure differences due to intron removal are impossible to control for.